

Query Match 24.2%; Score 89.6; DB 8; Length 456;
Best Local Similarity 67.2%; Pred. No. 6.3e-12;
Matches 125; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 31 TACAAACCTCTATACAGGATACAGCAATGTGAACACAGAGGTGATGATCTGAGGTTTC 90
DB 236 TGCTCCCCACAATATGAATTTGCCAGANTNGTTGCCAGAGGCCACACATGTCTGAGGTTTC 177
QY 91 ATCAAAACACTGGAGCCATGTTAGTTATTCCTTGTAGTAGTATGTCACACTCAACACC 150
DB 176 ATGAAGCACTGGAAACCATGTAGTTACTGTCTTGATAGTAGTGTGGGCTCAACACC 117
QY 151 AATGTCTCTCTGCTGTGATGGAGAAATGGGCTGCCGATTCACAGACTTCTTGATCTTTC 210
DB 116 AGTGCCCTCTGCTGTGATGAAGAAATGATGCCACTTGCCTGAAAGGATGTCAGCTG 57
QY 211 CTCAGC 216
DB 56 CCCAGC 51

RESULT 4
AV734452/c
LOCUS
DEFINITION AV734452 cda Homo sapiens CDNA clone cdaAOA01 5', mRNA sequence.
ACCESSION AV734452
VERSION AV734452.1 GI:10851997
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Yang, Y., Song, H., Peng, Y., Gu, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z. and Han, Z.
TITLE Homo sapiens CDNA cda clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhongjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
Location/Qualifiers
1..456
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="cdaAOA01"
/tissue_type="pneochromocytoma"
/dev_stage="Adult"
/lab_host="BM25.8"
/clone_lib="cda"
/note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"

Query Match 23.6%; Score 87.4; DB 1; Length 456;
Best Local Similarity 69.8%; Pred. No. 2.3e-11;
Matches 118; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 202 TGATCTTTCTCAGCATATGATCTCTCATATGAAACTCTCTCTGCTAGGAGTGCA 261
DB 399 TAATATCTTTAAGCATGACATGAATAAGTAATAATGAATTAAGCAGCTAGCATGTA 340
QY 262 GCCTACATCGTAAATCCAGCACCTTGGGAAGCTGAGGAGGAGGATTAATCGAGCCAG 321
DB 339 GCTCACATCTGCAATCCAGCACCTTGGGAGGCTGAGGAGGATGATGCTTGGAGCCAG 280

QY 322 GAGTTTGAGACCGGCTGGGCAATACAGCGAGACTCTCTCTCTTAAAAA 370
DB 279 GAGTTCAAGACCGGCTGGGCAACATAGTGAGACCCCATCTCCAAAAA 231

RESULT 5
BG912371/c
LOCUS
DEFINITION BG912371 744 bp mRNA linear EST 05-JUN-2001
602808438F1 NCI CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4940673
5', mRNA sequence.

ACCESSION BG912371
VERSION BG912371.1 GI:14292847
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1 (bases 1 to 744)
TITLE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov

Tissue Procurement: David N. Louis, M.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10879 row: k column: 10
High quality sequence stop: 711.
Location/Qualifiers
1..744
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4940673"
/tissue_type="anaplastic oligodendroglioma with 1p/19q
loss"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Brn67"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.3 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

ORIGIN
Query Match 23.4%; Score 86.6; DB 2; Length 744;
Best Local Similarity 80.8%; Pred. No. 3.7e-11;
Matches 101; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 246 CTGGCTAGGAGTGGCAGCCTACATCGCTAATCCAGCACCTTGGGAAGCTGAGGAGAG 305
DB 485 CTGGCTAGGCGTGGTGGCCACATAAGTAATACAGCAGCACTTTGGGAGCCAAAGTCAGGAG 426
QY 306 GATTACTGGAGGCCAGGAGTTTGAGCCAGGCTGGGCAATACAGGAGACTCTCTCTCT 365
DB 425 GATCACTTGAGCCCGAGGAGTTTGAGCAAGCCTGGGCAACATAGTGAGACTCCATCTCT 366
QY 366 AAAAA 370
DB 365 AAAAA 361

RESULT 6
AUI56956/c
LOCUS
DEFINITION AUI56956 PLACE1 Homo sapiens cDNA clone PLACE1005453 3', mRNA
sequence.
ACCESSION AUI56956


```
Best Local Similarity 65.4%; Pred. No. 7.3e-11;
Matches 125; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 180 GCTGCCGATTCACAGACTTCTTGATCTTTCTCCACGATATGATCTCTCATATGAACT 239
Db 355 GCTTTCTTTGCACTCTTTTCATAGGGGAATTCAGCAAAATTAATCATATAAAATG 296
QY 240 CTCCTCTCGCTAGGAGTGCAGCTACATCGCTAATCCAGCACCTTGGGAAGCTGAGG 299
Db 295 CCACTAGGGCTGGTGGGTGCTCAAGCTGTATCCAGTACTTTGGGAGGTAGAGG 236
QY 300 CAGAGGATTAATCTGGAGCCAGAGTTTGAGACAGGCTGGGCAATACAGCGAGACTCTC 359
Db 235 CAGGAGGATCAGTTGAGCCAGGAGTTTGAGATGAGCTGGGCAACACAGCAAGACCTG 176
QY 360 TCTCCTTAAAAA 370
Db 175 TCTCTACAAAA 165

RESULT 13
CK822235/c
LOCUS
DEFINITION
  ig80g07.y5 HR85 islet Homo sapiens cDNA clone IMAGE:5597749 5',
  mRNA sequence.
CK822235
CK822235.1 GI:44839160
EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Hominidae; Homo.
  1 (bases 1 to 537)
  Melton, D., Meadows, A., Clifton, S., Hillier, L., Marra, M., Pape, D.,
  Wylie, T., Martin, J., Blistain, A., Schmitt, A., Theising, B.,
  Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M.,
  McCann, R., Cole, R., Teagareishvili, R., Williams, T., Jackson, Y. and
  Bowers, Y.
  WashU-Harvard Pancreas EST Project
  Unpublished (2000)
  Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
  Endocrine Pancreas Consortium
  Harvard University, Howard Hughes Medical Institute
  Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
  MA 02138
  Tel: 617-495-1812
  Fax: 617-495-8557
  Email: dmelton@biohp.harvard.edu
  This read is a 5' RESEQUENCE of a previously sequenced pancreas
  clone
  Good hit to opposite strand read. . wrong orientation BUT PASSED FOR
  MOUSE-PANCREAS VERIFICATION
  Seq primer: -40UP from Gibco
  High quality sequence stop: 537.
  Location/Qualifiers
    1..537
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:5597749"
      /tissue_type="Purified pancreatic islet"
      /lab_host="DH10B"
      /clone_lib="HR85 islet"
      /note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
      NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
      Size-selected on agarose gel. Average insert size ~1kb. 5'
      XhoI site was destroyed after directional cloning.
      Amplified once. Contact information: Hiroshi Inoue, MD,
      Metabolism Div. (Alan Permutt Lab), Washington University
      School of Medicine, Box 8127, 660 South Euclid Ave., St.
      Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
      314-362-1916, Fax: 314-747-2692."
```

```
ORIGIN
  Query Match 23.1%; Score 85.4; DB 7; Length 537;
  Best Local Similarity 65.4%; Pred. No. 7.3e-11;
  Matches 125; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 180 GCTGCCGATTCACAGACTTCTTGATCTTTCTCCACGATATGATCTCTCATATGAACT 239
Db 316 GCTTTCTTTGCACTCTTTTCATAGGGGAATTCAGCAAAATTAATCATATAAAATG 257
QY 240 CTCCTCTCGCTAGGAGTGCAGCTACATCGCTAATCCAGCACCTTGGGAAGCTGAGG 299
Db 256 CCACTAGGGCTGGTGGGTGCTCAAGCTGTATCCAGTACTTTGGGAGGTAGAGG 197
QY 300 CAGGAGGATTAATCTGGAGCCAGAGTTTGAGACAGGCTGGGCAATACAGCGAGACTCTC 359
Db 196 CAGGAGGATCAGTTGAGCCAGGAGTTTGAGATGAGCTGGGCAACACAGCAAGACCTG 137
QY 360 TCTCCTTAAAAA 370
Db 136 TCTCTACAAAA 126

RESULT 14
AQ011703/c
LOCUS
DEFINITION
  HS 2172 A2 MF E07 CIT Approved Human Genomic Sperm Library D Homo
  sapiens genomic clone Plate=2172 Col=14 Row=I, genomic survey
  sequence.
ACCESSION
  AQ011703
VERSION
  AQ011703.1 GI:3171379
KEYWORDS
  GSS.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Hominidae; Homo.
  1 (bases 1 to 410)
  Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
  Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
  Hood, L.
  Sequence-tagged connectors: A sequence approach to mapping and
  scanning the human genome
  Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  10449764
  Contact: Mahairas GG, Wallace JC, Hood L
  High Throughput Sequencing Center
  University of Washington
  401 Queen Anne Avenue North, Seattle, WA 98109, USA
  Tel: (206) 616-3618
  Fax: (206) 616-3887
  Email: jwallace@u.washington.edu
  Sequence Tagged Connector
  Plate: 2172 row: I column: 14
  Class: BAC ends
  High quality sequence stop: 410.
  Location/Qualifiers
    1..410
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /clone="Plate=2172 Col=14 Row=I"
      /sex="male"
      /clone_lib="CIT Approved Human Genomic Sperm Library D"
      /note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in
      E-Coli DH10B"

ORIGIN
  Query Match 22.9%; Score 84.8; DB 9; Length 410;
  Best Local Similarity 65.0%; Pred. No. 1e-10;
  Matches 116; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 203 GATCTTTCTCAGCATATGATCTCTCATATGAAACTCTCTCTGGCTAGGAGTGGCAG 262
```

```
Db      266 GC|T|T|T|T|T|G|G|C|A|T|A|T|G|T|T|C|A|G|T|C|C|G|G|T|T|A|A|T|T|G|T|A|A|C|T|A|G|C|T|G|G|T|G|A|G|G|T|G| 207
Qy      263 C|T|A|C|A|T|C|G|T|A|T|C|C|A|G|C|A|C|T|T|G|G|A|A|G|C|T|G|A|G|C|A|G|A|T|T|A|C|T|G|G|A|G|G|C|A|G| 322
Db      206 C|T|C|A|C|C|C|T|G|T|A|T|C|C|A|G|C|G|C|T|T|G|G|A|G|C|T|G|A|G|C|A|G|T|A|T|T|G|T|T|G|A|G|C|T|C|A|G| 147
Qy      323 A|G|T|T|T|G|A|G|C|A|G|C|T|G|G|C|A|A|T|A|C|A|G|C|A|G|A|C|T|C|T|C|T|C|T|A|A|A|A| 370
Db      146 G|G|T|T|C|A|G|A|C|C|A|G|C|T|G|G|C|A|A|T|A|A|G|T|G|A|G|A|C|T|C|T|G|T|C|T|C|A|A|A|A| 99
```

RESULT 15

B92454/c

LOCUS

DEFINITION B92454 459 bp DNA linear GSS 25-JUN-1998
CIT-HSP-2172114.TF CIT-HSP Homo sapiens genomic clone 2172114,
genomic survey sequence.

ACCESSION

B92454

VERSION

B92454.1

KEYWORDS

GSS.

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominiidae; Homo.

1. (bases 1 to 459)

Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,

Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,

Simon,M. and Venter,J.C.

Use of a random BAC End Sequence Database for Sequence-Ready Map

Building (1998)

Unpublished (1998)

Other_GSSs: CIT-HSP-2172114.TR

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html

Seq primer: M13-21;

Class: BAC ends.

FEATURES

Source

1. .459

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="GDB:7103880"

/db_xref="taxon:9606"

/clone="2172114"

/sex="Male"

/cell_type="Sperm"

/clone_lib="CIT-HSP"

/notes="Vector: pBelosBAC11; Site_1: HindIII; Site_2:

HindIII"

ORIGIN

Query Match 22.9%; Score 84.6; DB 9; Length 459;

Best Local Similarity 69.0%; Pred. No. 1e-10;

Matches 116; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 203 G|A|T|C|T|T|C|T|C|A|G|A|T|A|T|A|C|T|C|T|C|A|T|A|T|G|A|A|A|C|T|C|T|C|T|G|G|C|T|A|G|A|G|T|G|G|C|A|G| 262

Db 275 G|C|T|T|T|T|T|G|G|A|C|A|T|A|T|G|T|T|C|A|G|T|C|C|G|G|T|A|A|A|T|T|G|T|A|A|C|T|A|G|C|T|G|G|T|G|A|G|T|G| 216

Qy 263 C|T|A|C|A|T|C|G|T|A|T|C|C|A|G|C|A|C|T|T|G|G|A|G|C|T|G|A|G|C|A|G|A|T|T|A|C|T|G|A|G|G|C|A|G| 322

Db 215 C|T|C|A|C|C|T|G|T|A|T|C|C|A|G|C|G|C|T|T|G|G|A|G|C|T|G|A|G|C|A|G|T|A|T|T|G|T|T|G|A|G|C|T|C|A|G| 156

Qy 323 A|G|T|T|T|G|A|G|A|C|C|A|G|C|T|G|G|C|A|A|T|A|C|A|G|C|A|G|A|C|T|C|T|C|T|C|T|A|A|A|A| 370

Db 155 G|G|T|T|C|C|A|G|A|C|C|A|G|C|C|T|G|C|G|C|A|A|T|A|A|A|G|T|G|A|G|A|C|T|C|T|G|T|C|T|C|A|A|A|A| 108

Search completed: December 7, 2005, 09:55:41
Job time : 4226.37 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 751.729 Seconds
(without alignments)
6427.437 Million cell updates/sec

Title: US-09-980-046B-10
Perfect score: 85
Sequence: 1 cggtagtgatctcttcgc.....tgcccactaagtagaaaaa 85

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_in.*
- 3: gb_env.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pr.*
- 9: gb_ro.*
- 10: gb_sts.*
- 11: gb_sy.*
- 12: gb_un.*
- 13: gb_vi.*
- 14: gb_htg.*
- 15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83.4	98.1	451	6	AX135336 Sequence
2	83.4	98.1	1318	8	BC001114 Homo sapi
3	83.4	98.1	3901	8	L11045 Human LMP7
4	83.4	98.1	19936	8	BX682530 Human DNA
5	83.4	98.1	19937	8	CR749765 Human DNA
6	83.4	98.1	35167	8	CT009502 Human DNA
7	83.4	98.1	47546	8	CR753889 Human DNA
8	83.4	98.1	63179	8	AL935043 Human DNA
9	83.4	98.1	66109	6	AX330551 Sequence
10	83.4	98.1	66109	6	AX411124 Sequence
11	83.4	98.1	66109	8	HSMHCAPG
12	83.4	98.1	95554	8	CR762476 Human DNA
13	83.4	98.1	110856	8	BX927138 Human DNA
14	83.4	98.1	114576	8	BX088556 Human DNA
15	83.4	98.1	116859	14	CR786563 Human DNA
16	83.4	98.1	126988	8	AL669918 Human DNA
17	83.4	98.1	166803	8	AL671681 Human DNA
18	83.4	98.1	177571	14	AC022567 Homo sapi

19	83.4	98.1	198285	6	CQ861705 Sequence
20	83.4	98.1	198285	6	AX411170 Sequence
21	83.4	98.1	198285	8	HSEVMHC
22	83.4	98.1	349980	6	CS039424 Sequence
23	82.4	96.9	9362	8	AY559263 Pan trogl
24	82.4	96.9	9362	8	AY559271 Pan trogl
25	82.4	96.9	9362	8	AY559277 Pan trogl
26	82.4	96.9	9362	8	AY559291 Pan trogl
27	82.4	96.9	9362	8	AY559293 Pan trogl
28	82.4	96.9	9362	8	AY559299 Pan trogl
29	82.4	96.9	9363	8	AY559261 Pan trogl
30	82.4	96.9	9366	8	AY559253 Pan trogl
31	82.4	96.9	9366	8	AY559257 Pan trogl
32	82.4	96.9	9366	8	AY559259 Pan trogl
33	82.4	96.9	9366	8	AY559264 Pan trogl
34	82.4	96.9	9366	8	AY559267 Pan trogl
35	82.4	96.9	9366	8	AY559275 Pan trogl
36	82.4	96.9	9366	8	AY559279 Pan trogl
37	82.4	96.9	9366	8	AY559280 Pan trogl
38	82.4	96.9	9366	8	AY559281 Pan trogl
39	82.4	96.9	9366	8	AY559285 Pan trogl
40	82.4	96.9	9366	8	AY559295 Pan trogl
41	82.4	96.9	9366	8	AY559297 Pan trogl
42	82.4	96.9	9367	8	AY559273 Pan trogl
43	82.4	96.9	9367	8	AY559289 Pan trogl
44	82.4	96.9	9368	8	AY559260 Pan trogl
45	82.4	96.9	9368	8	AY559262 Pan trogl

ALIGNMENTS

RESULT 1
AX135336
LOCUS AX135336 451 bp DNA linear PAT 29-MAY-2001
DEFINITION Sequence 144 from Patent WO0132928.
ACCESSION AX135336
VERSION AX135336.1 GI:14271685
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Farr S.
TITLE Methods of determining individual hypersensitivity to an agent
JOURNAL Patent: WO 0132928-A 144 10-MAY-2001;
FEATURES Phase-1 Molecular Toxicology Inc. (US)
Location/Qualifiers
source 1. .451
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 98.1%; Score 83.4; DB 6; Length 451;
Best Local Similarity 98.8%; Pred. No. 5.8e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGATCTTCGTGTTCTCTATTGAACAGCATTTCCCGGAGGAGTTT 60
Db 353 CGGTGAAGTGATCTTCGTGTTCTCTATTGAACAGCATTTCCCGGAGGAGTTT 412

Qy 61 CTGGGTGCCCCACTAAGTAGAGAAA 85
Db 413 CTGGGTGCCCCACTAAGTAGAGATAA 437

RESULT 2
BC001114
LOCUS BC001114 1318 bp mRNA linear PRI 28-JUL-2005
DEFINITION Homo sapiens proteasome (prosome, macropain) subunit, beta type, 8

(large multifunctional protease 7), transcript variant 1, mRNA
(CDNA clone MGC:1491 IMAGE:2967119), complete cds.

BC001114
VERSION BC001114.2 GI:33876008
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A.C., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzinski, M.I., Skalska, U., Smalusz, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1318)

NITH MGC Project
Direct Submission
Submitted (11-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NITH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:12654558.
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: ATCC/DCTB/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amad@systemsbio.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 3 Row: m Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 34335277.
Location/Qualifiers
1. .1318
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:1491 IMAGE:2967119"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NITH MGC 20"
/lab_host="DH10B-R"
/notes="vector: pOTB7"
1. .1318
/genes="PSMB8"
/notes="synonyms: RING10, MGC1491, D6S216, D6S216E"
/db_xref="GeneID:5696"
/db_xref="MIM:177046"

gene

ACCESSION BC001114
VERSION BC001114.2 GI:33876008
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A.C., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzinski, M.I., Skalska, U., Smalusz, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1318)

NITH MGC Project
Direct Submission
Submitted (11-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NITH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:12654558.
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: ATCC/DCTB/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amad@systemsbio.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 3 Row: m Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 34335277.
Location/Qualifiers
1. .1318
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:1491 IMAGE:2967119"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NITH MGC 20"
/lab_host="DH10B-R"
/notes="vector: pOTB7"
1. .1318
/genes="PSMB8"
/notes="synonyms: RING10, MGC1491, D6S216, D6S216E"
/db_xref="GeneID:5696"
/db_xref="MIM:177046"

gene

CDs

239. .1057
/genes="PSMB8"
/codon_start=1
/product="proteasome beta 8 subunit, isoform E1
protein"
/protein_id="AAH01114.1"
/db_xref="GI:12654559"
/db_xref="GeneID:5696"
/db_xref="MIM:177046"
/translation="MLIGTTPRDTTPSSWLTSLLVEAAPLDDTTLTPVSSGCGPL
EPTFQSLGGDGRNVQIEMAHGTTTLAKFQHGVIIVADSRASAGSYISALRVNVK
IINPVLCTMSGCAADCOYWERLLAKECRLAYLRNGERISVSAASKLLSNMCOYRG
MELSGSMICGDKKPGGLYVDEHGTRELRLSGNMFSGNCTYAYGVWDGSGYRNLSP
EAYDLGRRATATATHRDSYSGGVVNIHMKEDGVKVESTIDVSDLLHQREANQ"

ORIGIN

Query Match 98.1%; Score 83.4; DB 8; Length 1318;
Best Local Similarity 98.8%; Pred. No. 5.8e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCTCGGTCTTCTCTATTGTAACAAGCATTTCCCCAGGGAAGTTT 60
Db 1196 CGGTGAAGTGCATCTTCTCGGTCTTCTCTATTGTAACAAGCATTTCCCCAGGGAAGTTT 1255

QY 61 CTGGTGCCCCACTAAGTAGAAAA 85
Db 1255 CTGGTGCCCCACTAAGTAGATAA 1280

RESULT 3
HUMLMP7A
LOCUS HUMLMP7A
DEFINITION Human LMP7 gene, exons 1-7.
ACCESSION L11045
VERSION L11045.1 GI:307133
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 3901)
AUTHORS Meinhardt, T., Graf, U. and Hammerling, G.J.
TITLE Different genomic structure of mouse and human Lmp7 genes:
characterization of MHC-encoded proteasome genes
JOURNAL Immunogenetics 38 (5), 373-379 (1993)
PUBMED 8344725
COMMENT Original source text: Homo sapiens DNA.
FEATURES
source
1. .3901
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/cell_line="B lymphoblastoid cell line MANN"
1. .3901
/genes="LMP7"
precursor_RNA
1. .3901
/genes="LMP7"
mRNA
join(1. .355, 610. .784, 1545. .1692, 1851. .1962, 2368. .2497,
2893. .3097, 3578. .3901)
/genes="LMP7"
1. .355
/genes="LMP7"
exon
1. .355
/number=1
356. .609
/genes="LMP7"
intron
/number=1
610. .784
/genes="LMP7"
exon
785. .1544
/number=2
/genes="LMP7"
intron
/number=2

1545..1692
/gene="LMP7"
/number=3
1693..1850
/gene="LMP7"
/number=3
1851..1962
/gene="LMP7"
/number=4
1963..2367
/gene="LMP7"
/number=4
2368..2497
/gene="LMP7"
/number=5
2498..2892
/gene="LMP7"
/number=5
2893..3097
/gene="LMP7"
/number=6
3098..3577
/gene="LMP7"
/number=6
3578..3901
/gene="LMP7"
/number=7
3885..3890
/gene="LMP7"

polyA_site

ORIGIN

Query Match 98.1%; Score 83.4; DB 8; Length 3901;
Best Local Similarity 98.8%; Pred. No. 5.7e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTCATCTTCGCGTGTCTCTATTGAACAAGCATTTCCCCAGGGAAGTTT 60
Db 3805 CGGTGAAGTCATCTTCGCGTGTCTCTATTGAACAAGCATTTCCCCAGGGAAGTTT 3864

Qy 61 CTGGGTGCCCACTAAGTAGAATAA 85
Db 3865 CTGGGTGCCCACTAAGTAGAATAA 3889

RESULT 4
BX682530/c
LOCUS BX682530 19936 bp DNA linear PRI 18-MAY-2005
DEFINITION Human DNA sequence from clone DASS-193J13 on chromosome 6, complete sequence.
ACCESSION BX682530
VERSION BX682530.4 GI:42475735
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 19936)
Palmer,S.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Feb 7, 2004 this sequence version replaced gi:40067399.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6/MHC

DASS-193J13 is from a DNA-arts SSTO human bac library VECTOR:
pBelOBAC11

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: vegas@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

SOURCE

1..19936
/location=Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="DASS-193J13"
/clone_lib="DNA-arts-BAC.1-SSTO.1"

misc_feature

2000
/note="Clone_right_end: DASS-25606"
complement(join(7236..7759,8125..8261,8655..8814,
8992..9165,9343..9531,11058..11186,11352..11549,
13879..14084,14368..14498,16262..16376,16466..16962,
17374..17465))
/locus_tag="DASS-193J13.1-001"

mRNA

complement(join(7236..7759,8125..8261,8655..8814,
8992..9165,9343..9531,11058..11186,11352..11549,
13879..14084,14368..14498,16262..16376,16466..16962,
17374..17465))
/locus_tag="DASS-193J13.1-001"
/notes="match: ESTs: Em:903732.1 Em:AUI32766.1
Em:BI84592.1 Em:BM561472.1 Em:BQ277877.1
match: CDNAS: Em:BC002751.2 Em:M74447.1 Em:M84748.1
Em:Z22936.1"

CDS

complement(join(7580..7759,8125..8261,8655..8814,
8992..9165,9343..9531,11058..11186,11352..11549,
13879..14084,14368..14498,16262..16376,16466..16958))
/locus_tag="DASS-193J13.1-001"
/standard_name="OTTHUMP00000038912"
/codon_start=1
/protein_id="CAI41935.1"
/db_xref="GI:57209644"
/db_xref="GOA:Q5HY71"
/db_xref="InterPro:IPR001140"
/db_xref="InterPro:IPR003439"
/db_xref="InterPro:IPR003593"
/db_xref="InterPro:IPR005293"
/db_xref="UniProt/Tr:EMBL:Q5HY71"

/translation="MRPLDRLPWTSLLLVDALLLWQLQGLTLLPGLPGLWLEGTLL
LVGGLWLLKRLGLLGVFTLLPLCLATPLTVSLRALVAGASRPDPARVASAPWSWL
LVGGAGLWSLWAVLSPGAEQEQDQNNKVMWLLKSLRDLPLVLAAPFLV
NALVGLTLPYSGRVIDILGDDPHAFASAIFFMCLFSFSGSLSAGCGGCTYTM
SRNLRIEOLFFSLRLDQFTQTKTGLNSRLSDTTTMSNWLPLNANVLLRSLV
KVGLYGFMLSIPLRLTLLSLHMPPTIAAEKVYNTRHQEVLEIQAARAGQVVR
AVGLQTVRFSGAEHEVCYKEALEOCRLQYWRDLERALLVRRVHLGVQMLL
SCGLQQQDGLTQGSLLSEMIYQSSVSVYQTLVIYIGDMLSVGAARVFSIMDRQ
KPSFPGTLPNTYQGVVQDFSPAYPNRPDLVRLKGLTFLRPGVETLVAPNGSG
KSTVAALLQNTYQGVVLDLDEKPIQYEHCVLHSQVSVGVQVLFSGSVRNIIAY
GLOSCDDKYMARAQAHAADDFTQEMEHGIYTDVGEKGSOLAAGQKORLALARLV
PRVILDEATSDLVQCEQALQDWNRSGRDRLVIAHRLQVQAQHLVLQEGKQK
LAQEGQDLYSLRVQRLMD"
join(complement(17374..17465),complement(16466..16962),
complement(16262..16376),complement(14368..14498),
complement(11058..11186),complement(11352..11549),
complement(11058..11186),complement(9343..9531),
complement(8992..9165),complement(8655..8814),
complement(8125..8261),
complement(BX296564.7:73552..74037))

gene

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 63179)
Tracey,A.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Oct 30, 2002 this sequence version replaced gi:24415778.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6/MHC
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
DAQB-69D7 is from a DNA-arts QBL human bac library VECTOR: pBelOBAC11.
FEATURES
source
1. .63179
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="DAQB-69D7"
/clone_lib="DNA-arts-BAC.1-QBL.1"
misc_feature
1
/note="Clone left end: DAQB-69D7"
complement(join(1711..2196,9277..9413,9807..9966,10144..10317,10495..10683,12210..12338,12504..12701,15031..15236,15520..15650,17414..17528,17618..18114,18532..18623))
/gene="TAP2"
/locus_tag="DAQB-69D7.2-002"
complement(join(1711..2196,9277..9413,9807..9966,10144..10317,10495..10683,12210..12338,12504..12701,15031..15236,15520..15650,17414..17528,17618..18114,18532..18623))
/gene="TAP2"
/locus_tag="DAQB-69D7.2-002"
complement(join(1711..2196,9277..9413,9807..9966,10144..10317,10495..10683,12210..12338,12504..12701,15031..15236,15520..15650,17414..17528,17618..18114,18532..18623))
/gene="TAP2"
/locus_tag="DAQB-69D7.2-002"
/product="transporter 2, ATP-binding cassette, sub-family B (MDR/TAP)"
B (MDR/TAP)
/note="match: cDNAs: AF078671.1 AF105151.1"
complement(1723..1728)
/gene="TAP2"
/locus_tag="DAQB-69D7.2-002"
complement(join(2167..2196,9277..9413,9807..9966,10144..10317,10495..10683,12210..12338,12504..12701,15031..15236,15520..15650,17414..17528,17618..18110))
/gene="TAP2"
/locus_tag="DAQB-69D7.2-002"
/standard_name="OTTHUMP0000014665"
/note="match: cDNAs: O95410 P36372 Q03519 Q28438"
/codon_start=1
/product="transporter 2, ATP-binding cassette, sub-family B (MDR/TAP)"
B (MDR/TAP)
/protein_id="CA118621.1"

/db_xref="GI:56207306"
/db_xref="GGA:Q9UP03"
/db_xref="InterPro:IPR001140"
/db_xref="InterPro:IPR003439"
/db_xref="InterPro:IPR003593"
/db_xref="InterPro:IPR005293"
/db_xref="InterPro:IPR011527"
/db_xref="UniProt/TrEMBL:Q9UP03"
/translation="MRLPDLRPWTSLLVDAALLWLLQGLTLLPQGLPGLMEGLRLVGYAGLSWSLWAVLSPGAEQEDQNNKVMRLIKLSRDLPLVAFAFFVLAVLGETLIPHYSGRVIDILGDFDPHAFASAIFFMCLFSFGSLGACGCGCTYTM SRINLRIRQLFSSLLRQDLGFFQETKTGELNSRLSDTTLMSNMLPLANVLLRSIVKVGLYGMFSISPRITLLSLHMPPTIAAEKYVNTRHQVLRLEIQAVARAGQVREAVGGLTVTSFGAEEHEVCYKEALQCRQLYWRDLERALLYLVRVHLGVQMLMLSCGLQMQGLTQGSLLSFMVYQESVSVQTLVIYIGDMLSNVGAEEKVFSYMDRQPNLPSGTLTAPTTLQGVVRFQDVSPAYPNRPDRPVKGLTFTLRPEVTALVPGNSGKSTVAALLQNYOPTGGVLLDEKPIQSQVHCYVLSQVSVGQEPVLFSGSVRNNTAYGLQSCDDKVMAAAQAHAADDFIOEMEHGIYTDVGEKGSQAAGQKQRLAIALVRDPRVLLDEATSLDVQCEQAQTLWKPMIF"
complement(join(8366..8911,9277..9413,9807..9966,10144..10317,10495..10683,12210..12338,12504..12701,15031..15236,15520..15650,17414..17528,17618..18114,18532..18623))
/gene="TAP2"
/locus_tag="DAQB-69D7.2-001"
complement(join(8366..8911,9277..9413,9807..9966,10144..10317,10495..10683,12210..12338,12504..12701,15031..15236,15520..15650,17414..17528,17618..18114,18532..18623))
/gene="TAP2"
/locus_tag="DAQB-69D7.2-001"
/product="transporter 2, ATP-binding cassette, sub-family B (MDR/TAP)"
B (MDR/TAP)
/note="match: cDNAs: BC002751.1 M74447.1 M84748.1 222936.1"
complement(8366)
/gene="TAP2"
/locus_tag="DAQB-69D7.2-001"
complement(join(8783..8911,9277..9413,9807..9966,10144..10317,10495..10683,12210..12338,12504..12701,15031..15236,15520..15650,17414..17528,17618..18110))
/gene="TAP2"
/locus_tag="DAQB-69D7.2-001"
/standard_name="OTTHUMP0000014664"
/note="match: proteins: O95410 P36372 Q03519 Q28438"
/codon_start=1
/product="transporter 2, ATP-binding cassette, sub-family B (MDR/TAP)"
B (MDR/TAP)
/protein_id="CA118622.1"
/db_xref="GI:56207307"
/db_xref="InterPro:IPR001140"
/db_xref="InterPro:IPR003439"
/db_xref="InterPro:IPR003593"
/db_xref="InterPro:IPR005293"
/translation="MRLPDLRPWTSLLVDAALLWLLQGLTLLPQGLPGLMEGLRLVGYAGLSWSLWAVLSPGAEQEDQNNKVMRLIKLSRDLPLVAFAFFVLAVLGETLIPHYSGRVIDILGDFDPHAFASAIFFMCLFSFGSLGACGCGCTYTM SRINLRIRQLFSSLLRQDLGFFQETKTGELNSRLSDTTLMSNMLPLANVLLRSIVKVGLYGMFSISPRITLLSLHMPPTIAAEKYVNTRHQVLRLEIQAVARAGQVREAVGGLTVTSFGAEEHEVCYKEALQCRQLYWRDLERALLYLVRVHLGVQMLMLSCGLQMQGLTQGSLLSFMVYQESVSVQTLVIYIGDMLSNVGAEEKVFSYMDRQPNLPSGTLTAPTTLQGVVRFQDVSPAYPNRPDRPVKGLTFTLRPEVTALVPGNSGKSTVAALLQNYOPTGGVLLDEKPIQSQVHCYVLSQVSVGQEPVLFSGSVRNNTAYGLQSCDDKVMAAAQAHAADDFIOEMEHGIYTDVGEKGSQAAGQKQRLAIALVRDPRVLLDEATSLDVQCEQAQTLWKPMIF"
complement(join(20595..20925,21409..21613,22012..22141,22550..22661,22820..22967,24157..24581))
/gene="P5MB8"
/locus_tag="DAQB-69D7.6-001"
complement(join(20595..20925,21409..21613,22012..22141,

```
22550. .22661,22820. .22967,24157. .24581))
/gene="PSMB8"
/locus tag="DAQB-69D7.6-001"
/product="proteasome (prosome, macropain) subunit, beta
type, 8 (large multifunctional protease 7)"
/notes="match: ESTs: BM547457.1 BU935203.1
match: cDNAs: BC001114.1 X62598.1"
complement(join(20595. .20925,21409. .21613,22012. .22141,
22550. .22661,22820. .22967,23728. .23925))
/gene="PSMB8"
/locus tag="DAQB-69D7.6-002"
complement(join(20595. .20925,21409. .21613,22012. .22141,
22550. .22661,22820. .22967,23728. .23925))
/gene="PSMB8"
/locus tag="DAQB-69D7.6-002"
/product="proteasome (prosome, macropain) subunit, beta
type, 8 (large multifunctional protease 7)"
/notes="match: ESTs: BM547457.1
match: cDNAs: AK031801.1 U17496.1 U17497.1 U22033.1"
complement(20595)
/gene="PSMB8"
/locus tag="DAQB-69D7.6-002"
complement(20613. .20618)
/gene="PSMB8"
/locus tag="DAQB-69D7.6-002"
complement(join(20776. .20925,21409. .21613,22012. .22141,
22550. .22589,22820. .22967,23728. .23901))
/gene="PSMB8"
/locus tag="DAQB-69D7.6-003"
complement(join(20776. .20925,21409. .21613,22012. .22141,
22550. .22589,22820. .22967,23728. .23901))
/gene="PSMB8"
/locus tag="DAQB-69D7.6-003"
/product="proteasome (prosome, macropain) subunit, beta
type, 8 (large multifunctional protease 7)"
/notes="match: ESTs: BQ685477.1 BQ686592.1"
complement(join(20837. .20925,21409. .21613,22012. .22141,
22550. .22661,22820. .22967,24157. .24291))
/gene="PSMB8"

Query Match 98.1%; Score 83.4; DB 8; Length 63179;
Best Local Similarity 98.8%; Pred. No. 5.6e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 20698 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 20639

Qy 61 CTGGTGCCCCCACTAAGTAGAAAA 85
Db 20638 CTGGTGCCCCCACTAAGTAGAAAA 20614

RESULT 9
AX330551 66109 bp DNA linear PAT 09-JAN-2002
LOCUS
DEFINITION Sequence 1060 from Patent WO0194629.
ACCESSION AX330551
VERSION AX330551.1 GI:18103529
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrikan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
Gene sets
Patent: WO 0194629-A 1060 13-DEC-2001;
Avalon Pharmaceuticals (US)
Location/Qualifiers
```

```
source 1. .66109
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 98.1%; Score 83.4; DB 6; Length 66109;
Best Local Similarity 98.8%; Pred. No. 5.6e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 37840 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 37899

Qy 61 CTGGTGCCCCCACTAAGTAGAAAA 85
Db 37900 CTGGTGCCCCCACTAAGTAGAAAA 37924

RESULT 10
AX411124 66109 bp DNA linear PAT 14-JUN-2002
LOCUS
DEFINITION Sequence 3771 from Patent WO0229103.
ACCESSION AX411124
VERSION AX411124.1 GI:21443829
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
Gene expression profiles in liver cancer
Patent: WO 0229103-A 3771 11-APR-2002;
GENE LOGIC INC (US)
Location/Qualifiers
source 1. .66109
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/notes="EMBL/GenBank Accession No. X66401"

ORIGIN

Query Match 98.1%; Score 83.4; DB 6; Length 66109;
Best Local Similarity 98.8%; Pred. No. 5.6e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 37840 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 37899

Qy 61 CTGGTGCCCCCACTAAGTAGAAAA 85
Db 37900 CTGGTGCCCCCACTAAGTAGAAAA 37924

RESULT 11
HSMHCAPG 66109 bp DNA linear PRI 18-APR-2005
LOCUS
DEFINITION H.sapiens genes TAP1, TAP2, LMP2, LMP7 and DOB.
ACCESSION X66401 S57528
VERSION X66401.1 GI:34634
KEYWORDS DOB gene; LMP2 gene; LMP7 gene; major histocompatibility complex
class II; tap1 gene; tap2 gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Beck, S., Kelly, A., Radley, E., Khurshid, F., Alderton, R.P. and
Trowsdale, J.
```

TITLE DNA sequence analysis of 66 kb of the human MHC class II region encoding a cluster of genes for antigen processing

JOURNAL J. Mol. Biol. 228 (2), 433-441 (1992)

PUBMED 1453454

REFERENCE 2 (bases 1 to 66109)

AUTHORS Glynn, R., Kerr, L.A., Mockridge, I., Beck, S., Kelly, A. and Trowdale, J.

TITLE The major histocompatibility complex-encoded proteasome component LMP7: alternative first exons and post-translational processing

JOURNAL Eur. J. Immunol. 23 (4), 860-866 (1993)

PUBMED 8458375

REFERENCE 3 (bases 1 to 66109)

AUTHORS Beck, S.

TITLE Direct Submission

JOURNAL Submitted (21-MAY-1992) S. Beck, Imperial Cancer Research Fund, ICRF, 44 Lincoln's Inn Fields, London WC2A 3PX, UK

COMMENT On Sep 23, 2004 this sequence version replaced gi:301893.

FEATURES

source 1..66109

organism="Homo sapiens"

mol_type="Genomic DNA"

db_xref="taxon:9606"

chromosome="6"

misc_signal <1..>66109

notes="classII-J-box"

classII-X-box

classII-Y-box

1..272

/rpt_family="ALU"

2517..2811

/rpt_family="ALU"

2927..4135

/rpt_family="LINE"

4136..4864

/rpt_family="LINE"

4662..4864

/note="mer13"

/rpt_family="MER"

5182..5514

/rpt_family="ALU"

5525..5876

/rpt_family="LINE"

5672..5860

/note="mer13"

/rpt_family="MER"

6391..6689

/rpt_family="ALU"

8843..9091

/rpt_family="ALU"

13973..14135

/rpt_family="ALU"

14542..14795

/rpt_family="ALU"

14820..15595

/rpt_family="LINE"

15650..15789

/rpt_family="ALU"

15907..16194

/rpt_family="ALU"

16853..17150

/rpt_family="ALU"

18240..18529

/rpt_family="ALU"

complement(join(19127..19254,20154..20295,20525..20654,21265..21396,22455..22522,24371..24430))

/gene="LMP2"

complement(join(19127..19254,20154..20295,20525..20654,21265..21396,22455..22522,24371..24430))

/gene="LMP2"

function="antigen processing"

notes="proteasome related"

codon_start=1

protein_id="CAA47024.1"

exon

complement(19127..19254)

/gene="LMP2"

/number=6

complement(19253..20155)

/gene="LMP2"

/number=5

19588..19910

/rpt_family="ALU"

complement(20154..20295)

/gene="LMP2"

/number=5

complement(20294..20524)

/gene="LMP2"

/number=4

complement(20525..20654)

/gene="LMP2"

/number=4

complement(20653..21266)

/gene="LMP2"

/number=3

complement(21265..21396)

/gene="LMP2"

/number=3

complement(21395..22456)

/gene="LMP2"

/number=2

21438..21741

/rpt_family="ALU"

complement(22455..22522)

/gene="LMP2"

/number=2

complement(22521..24372)

/gene="LMP2"

/number=1

complement(24011..24707)

complement(24371..24430)

/gene="LMP2"

/number=1

24506..24519

24826..24838

/note="ISRE"

24896..24909

25021..25025

join(25024..25621,26158..26272,26421..26551,27511..27716,28143..28340,29542..29670,29820..30008,30568..30741,30985..31147,31456..31592,32875..33081)

/gene="TAP1"

join(25024..25621,26158..26272,26421..26551,27511..27716,28143..28340,29542..29670,29820..30008,30568..30741,30985..31147,31456..31592,32875..33081)

/gene="TAP1"

function="antigen processing"

/note="peptide pump"

codon_start=1

product="ABC-transporter"

protein_id="CAA47025.1"

/db_xref="GI:34636"

/db_xref="GOA:Q03518"

/db_xref="InterPro:IPR001140"

/db_xref="InterPro:IPR003439"

/db_xref="InterPro:IPR003593"

/db_xref="InterPro:IPR005293"

/db_xref="InterPro:IPR009078"

GC_signal

misc_signal

GC_signal

CAAT_signal

gene

CDS

```
/db_xref="UniProt/Swiss-Prot:Q03518"
/translators="MASSRCPAPRGRCCLFASLAWLGTVLLLLADWVLLRTALPRIF
SLVPTALPLLRVWAVGLSRWAVLWGLACGVRATVGTGSKSENAGACQWLAALPKLAAA
LGLALPGALFRELISWAPGSADSTRLLHMGSHPTAFVVSAAALPAALWHLGSL
WFGGQGGSNPVRRLGLGSGTBRSLFLVLVLSLGGEMAIPTFTGRLTDWILQD
GSADFTFRNLTLITIASAVLEFVGDGLYNNMTMGHVSHLQGEVFGAVROTEFF
QQQNTGIMSRTVDTLSLSDLSLFLVLYRGLCLLGLMLGWSVSLTNTLIT
LPLFLPKKVVWYQLLEVQVRESLAKSQVAIEALSAMPTVRSFANEGEAKPFE
KLOEIKTAVOGLAVAVNSWTTISGMLKVLGILYTGQLVTSGLVSSGNLVTFVLY
QMOFTQAVELLSTYPRQKAVGSEKIPEYLDRTPRCPGSLTPLHLGLVQFQDV
SFAYPNRPDLVLQGLFTLPRCEVTALVPGSGSKSTVAALLQNLVQPTGSGQLLDG
KPLPQYBHRVYLHQAAGQEPQVFGSLQENIAYGLTQKPTMBEITAAVKSANSP
ISGLPQGYDTEVDEAGSQLSGQORQAVALARLIRKPCVLILDDATSALDANSOLOVE
QLLYSPERYSRSVLLITQHLSLVEQADHILFLEGGAIREGGTHQQLMEKKGCYMWV
QAPADAPE"
25024..25621
/gene="TAP1"
/number=1
25622..26157
/gene="TAP1"
/number=1
26158..26272
/gene="TAP1"
/number=2
26273..26420
/gene="TAP1"
/number=2
26421..26551
/gene="TAP1"
/number=3
26552..27510
/gene="TAP1"
/number=3
26643..26752
/number=3
/feature="ca"
/rpt_unit="26643..26644"
26919..26967
/gene="TAP1"
/rpt_family="LAD"
27004..27064
repeat_region
98.1%; Score 83.4; DB 8; Length 66109;
Best Local Similarity 98.8%; Pred. No. 5.6e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 CGGTGAAGTGATCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 37840 CGGTGAAGTGATCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 37899

Qy 61 CTGGGTGCCCCACTAAGTAGAATAA 85
Db 37900 CTGGGTGCCCCACTAAGTAGAATAA 37924
```

```
RESULT 12
CR762476/c CR762476 95554 bp DNA linear PRI 19-OCT-2004
LOCUS Human DNA sequence from clone DAAP-57C1 on chromosome 6, complete
DEFINITION sequence.
ACCESSION CR762476
VERSION CR762476.2 GI:54114714
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 95554)
AUTHORS Wood, J.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
```

COMMENT

On Oct 13, 2004 this sequence version replaced gi:52000322.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6/MHC>
DAAP-57C1 is from a APD human bac - B Lymphoblastoid Cell Line
Library VECTOR: pBelobAC11.

FEATURES

Location/Qualifiers
1..95554
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="DAAP-57C1"
/clone_lib="APD"

ORIGIN

Query Match 98.1%; Score 83.4; DB 8; Length 95554;
Best Local Similarity 98.8%; Pred. No. 5.5e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 CGGTGAAGTGATCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 68596 CGGTGAAGTGATCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 68537
```

```
Qy 61 CTGGGTGCCCCACTAAGTAGAATAA 85
Db 68536 CTGGGTGCCCCACTAAGTAGAATAA 68512
```

RESULT 13

```
BA927138/c BA927138 110856 bp DNA linear PRI 10-MAR-2004
LOCUS Human DNA sequence from clone DAMA-91123 on chromosome 6, complete
DEFINITION sequence.
```

ACCESSION BA927138

VERSION BA927138.5 GI:45379240

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE 1 (bases 1 to 110856)

AUTHORS Sycamore, N.

TITLE Direct Submission

JOURNAL Submitted (09-MAR-2004) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

COMMENT
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Mar 11, 2004 this sequence version replaced gi:42820837.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr6/MHC
 DANA-91123 is from the DNA-Arts human BAC library MANN.1 VECTOR: pBelOBAC11.
 ----- Location/Qualifiers
 source
 1. .110856
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="6"
 /clone="DANA-91123"
 /clone_lib="DNA-Arts BAC library MANN.1"

 FEATURES
 misc_feature
 1
 /notes="Clone left_end: DASS-67H14"
 misc_feature
 114576
 /notes="Clone right_end: DASS-67H14"

 ORIGIN
 Query Match 98.1%; Score 83.4; DB 8; Length 110856;
 Best Local Similarity 98.8%; Pred. No. 5.5e-18;
 Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT 60
 |||||
 Db 8030 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT 7971
 |||||
 QY 61 CTGGTGCCCCCTAAGTAGAAAA 85
 |||||
 Db 7970 CTGGTGCCCCCTAAGTAGAATAA 7946
 |||||
 RESULT 14
 BX088556/c
 LOCUS BX088556 114576 bp DNA linear PRI 18-MAY-2005
 DEFINITION Human DNA sequence from clone DASS-67H14 on chromosome 6, complete sequence.
 ACCESSION BX088556
 VERSION BX088556.11 GI:32968209
 KEYWORDS HTG;
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 114576)
 Direct Submission
 Sycamore.N.
 Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
 Cambridge, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk

Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
 Clone requests: clonerequest@sanger.ac.uk
 On Jul 18, 2003 this sequence version replaced gi:32563203.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr6/MHC
 DASS-67H14 is from a DNA-Arts SSTS human bac library VECTOR: pBelOBAC11.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: vegas@sanger.ac.uk

 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr6/MHC
 DASS-67H14 is from a DNA-Arts SSTS human bac library VECTOR: pBelOBAC11.
 ----- Location/Qualifiers
 source
 1. .114576
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="6"
 /clone="DASS-67H14"
 /clone_lib="DNA-Arts-BAC.1-SSTS.1"

 misc_feature
 1
 /notes="Clone left_end: DASS-67H14"
 misc_feature
 114576
 /notes="Clone right_end: DASS-67H14"

 ORIGIN
 Query Match 98.1%; Score 83.4; DB 8; Length 114576;
 Best Local Similarity 98.8%; Pred. No. 5.5e-18;
 Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT 60
 |||||
 Db 1607 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT 1548
 |||||
 QY 61 CTGGTGCCCCCTAAGTAGAAAA 85
 |||||
 Db 1547 CTGGTGCCCCCTAAGTAGAATAA 1523
 |||||
 RESULT 15
 CR786563/c
 LOCUS CR786563 116859 bp DNA linear HTG 26-OCT-2004
 DEFINITION Homo sapiens chromosome 6 clone DADB-30110, 2 unordered pieces.
 ACCESSION CR786563
 VERSION CR786563.5 GI:54633101
 KEYWORDS HTG; HTGS PHASE1; HTGS_CANCELLED.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 116859)
 Direct Submission
 Sims.S.
 Submitted (22-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
 Cambridge, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Oct 25, 2004 this sequence version replaced gi:54306012.


```

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bDB30110
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 78856 bases at least Q40
Consensus quality: 78857 bases at least Q30
Consensus quality: 78857 bases at least Q20
Insert size: 116759; sum-of-contigs
Insert size: 84033; 2.3% error; agarose-fp
Quality coverage: 6.62x in Q20 bases; sum-of-contigs Quality
coverage: 9.20x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 78857: contig of 78857 bp in length
*
* 78858 78957: gap of 100 bp
*
* 78958 116859: contig of 37902 bp in length.
FEATURES
      source              1..116859
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="6"
                        /clone="DADB-30110"
                        /clone_lib="DNA-arts-BAC.1-DBB.1"
      misc_feature        1..78857
                        /note="assembly_fragment:01987"
      misc_feature        78958..116859
                        /note="assembly_fragment:02008"
ORIGIN
Query Match          98.1%; Score 83.4; DB 14; Length 116859;
Best Local Similarity 98.8%; Pred. No. 5.5e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
      |||
Db      105184 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 105125

Qy      61  CTGGGTGCCCTACTAGTAGAATAA 85
      |||
Db      105124 CTGGGTGCCCTACTAGTAGAATAA 105100

Search completed: December 7, 2005, 04:22:18
Job time : 753.729 secs

```

This Page Blank (100%)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:31:40 ; Search time 211.652 Seconds
(without alignments)
3320.997 Million cell updates/sec

Title: US-09-980-046B-10
Perfect score: 85
Sequence: 1 cggggaagtcacatctctgc.....tgccccactaagtagagaaaa 85

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA_Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83.4	98.1	1126	8	US-10-741-600-161
2	83.4	98.1	1143	3	US-09-925-302-79
3	83.4	98.1	1143	3	US-09-925-302-79
4	83.4	98.1	1271	5	US-10-084-817-272
5	83.4	98.1	1320	6	US-10-159-563-162
6	83.4	98.1	1539	8	US-10-741-600-162
7	83.4	98.1	2027	5	US-10-198-846-10394
8	83.4	98.1	3674	5	US-10-084-817-258
9	83.4	98.1	16175	8	US-10-741-600-17596
10	83.4	98.1	20773	8	US-10-741-600-17595
11	83.4	98.1	28789	8	US-10-741-600-17894
12	83.4	98.1	66109	3	US-09-880-107-3768
13	83.4	98.1	66109	3	US-09-873-367C-1060
14	83.4	98.1	66109	9	US-10-843-641A-1060
15	83.4	98.1	198285	8	US-09-880-107-3814
16	83.4	98.1	198285	8	US-10-775-169-338
17	83.4	98.1	198285	9	US-10-756-149-12
18	54.6	64.2	11398	6	US-10-311-455-2411
19	54.6	64.2	11398	6	US-10-311-455-2412
20	54.6	64.2	3673778	6	US-10-312-841-1
21	54.6	64.2	3673778	6	US-10-312-841-2
22	28.2	33.2	1530	4	US-09-925-065A-52319
23	28.2	33.2	1530	4	US-09-925-065A-52320

ALIGNMENTS

RESULT 1

US-10-741-600-161
; Sequence 161, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARCILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10741.600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 1126
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-161

Query Match 98.1%; Score 83.4; DB 8; Length 1126;
Best Local Similarity 98.8%; Pred. No. 3.8e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTGCTGTTCTCTATTTCGAACAAGCATTTCCCGGGAAGTTT 60
Db 1022 CGGTGAAGTGCATCTTCTGCTGTTCTCTATTTCGAACAAGCATTTCCCGGGAAGTTT 1081
Qy 61 CTGGGTGCCCCCACTAAGTAGAATAA 85
Db 1082 CTGGGTGCCCCCACTAAGTAGAATAA 1106

RESULT 2

US-09-925-302-79
; Sequence 79, Application US/0925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925.302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 1143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1126)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-302-79

Query Match 98.1%; Score 83.4; DB 3; Length 1143;
Best Local Similarity 98.8%; Pred. No. 3.8e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
DB 1006 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1065

QY 61 CTGGTGCCCCCACTAAGTAGAAAAA 85
DB 1066 CTGGTGCCCCCACTAAGTAGAAAAA 1090

RESULT 3
US-09-925-302-79
Sequence 79, Application US/09925302
Publication No. US20030064072A9
GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 1143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1126)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-302-79

Query Match 98.1%; Score 83.4; DB 3; Length 1143;
Best Local Similarity 98.8%; Pred. No. 3.8e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
DB 1006 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1065

QY 61 CTGGTGCCCCCACTAAGTAGAAAAA 85
DB 1066 CTGGTGCCCCCACTAAGTAGAAAAA 1090

RESULT 4
US-10-084-817-272
Sequence 272, Application US/10084817
Publication No. US20030119009A1
GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet

; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 272
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 2818482CB1
US-10-084-817-272

Query Match 98.1%; Score 83.4; DB 5; Length 1271;
Best Local Similarity 98.8%; Pred. No. 4e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
DB 1160 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1219

QY 61 CTGGTGCCCCCACTAAGTAGAAAAA 85
DB 1220 CTGGTGCCCCCACTAAGTAGAAAAA 1244

RESULT 5
US-10-159-563-162
Sequence 162, Application US/10159563
Publication No. US20040009154A1
GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
; FILE REFERENCE: 11613.56US11
; CURRENT APPLICATION NUMBER: US/10/159,563
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 10/133,937
; PRIOR FILING DATE: 2002-04-25
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 162
; LENGTH: 1320
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-159-563-162

Query Match 98.1%; Score 83.4; DB 6; Length 1320;
Best Local Similarity 98.8%; Pred. No. 4e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
DB 1198 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1257

QY 61 CTGGTGCCCCCACTAAGTAGAAAAA 85
DB 1258 CTGGTGCCCCCACTAAGTAGAAAAA 1282

RESULT 6
US-10-741-600-162
Sequence 162, Application US/10741600
Publication No. US20050026169A1
GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.

```
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741.600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 162
; LENGTH: 1539
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-741-600-162

Query Match          98.1%; Score 83.4; DB 8; Length 1539;
Best Local Similarity 98.8%; Pred. No. 4.2e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCACTCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
    |||||
Db 1435 CGGTGAAGTGCACTCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1494

Qy 61 CTGGGTGCCCCCACTAAGTAGAATAA 85
    |||||
Db 1495 CTGGGTGCCCCCACTAAGTAGAATAA 1519

RESULT 7
US-10-198-846-10394
; Sequence 10394, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10394
; LENGTH: 2027
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1804, 2025, 2026, 2027
; OTHER INFORMATION: n = A,T,C or G
; US-10-198-846-10394

Query Match          98.1%; Score 83.4; DB 5; Length 2027;
Best Local Similarity 98.8%; Pred. No. 4.6e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCACTCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
    |||||
Db 1282 CGGTGAAGTGCACTCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1341

Qy 61 CTGGGTGCCCCCACTAAGTAGAATAA 85
    |||||
Db 1342 CTGGGTGCCCCCACTAAGTAGAATAA 1366

RESULT 8
US-10-084-817-258
; Sequence 258, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
```

```
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 258
; LENGTH: 3674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 391741.16
; US-10-084-817-258

Query Match          98.1%; Score 83.4; DB 5; Length 3674;
Best Local Similarity 98.8%; Pred. No. 5.5e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCACTCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
    |||||
Db 3501 CGGTGAAGTGCACTCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 3560

Qy 61 CTGGGTGCCCCCACTAAGTAGAATAA 85
    |||||
Db 3561 CTGGGTGCCCCCACTAAGTAGAATAA 3585

RESULT 9
US-10-741-600-17596
; Sequence 17596, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741.600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17596
; LENGTH: 16175
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-741-600-17596

Query Match          98.1%; Score 83.4; DB 8; Length 16175;
Best Local Similarity 98.8%; Pred. No. 8.9e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCACTCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
    |||||
Db 10072 CGGTGAAGTGCACTCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 10131

Qy 61 CTGGGTGCCCCCACTAAGTAGAATAA 85
    |||||
Db 10132 CTGGGTGCCCCCACTAAGTAGAATAA 10156

RESULT 10
US-10-741-600-17595
; Sequence 17595, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
```

```
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741.600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17595
; LENGTH: 20773
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-741-600-17595

Query Match      98.1%; Score 83.4; DB 8; Length 20773;
Best Local Similarity 98.8%; Pred. No. 9.6e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 19163 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 19222

Qy 61 CTGGGTGCCCCACTAAGTAGAATAA 85
Db 19223 CTGGGTGCCCCACTAAGTAGAATAA 19247

RESULT 11
US-10-741-600-17894
; Sequence 17894, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741.600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17894
; LENGTH: 28789
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-741-600-17894

Query Match      98.1%; Score 83.4; DB 8; Length 28789;
Best Local Similarity 98.8%; Pred. No. 1.1e-19;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 3965 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 4024

Qy 61 CTGGGTGCCCCACTAAGTAGAATAA 85
Db 4025 CTGGGTGCCCCACTAAGTAGAATAA 4049

RESULT 12
US-09-880-107-3768
; Sequence 3768, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
```

```
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3768
; LENGTH: 66109
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(66109)
; OTHER INFORMATION: n=a,t,g or c
US-09-880-107-3768

Query Match      98.1%; Score 83.4; DB 3; Length 66109;
Best Local Similarity 98.8%; Pred. No. 1.4e-19;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 37840 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 37899

Qy 61 CTGGGTGCCCCACTAAGTAGAATAA 85
Db 37900 CTGGGTGCCCCACTAAGTAGAATAA 37924

RESULT 13
US-09-873-367C-1060
; Sequence 1060, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1060
; LENGTH: 66109
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(66109)
; OTHER INFORMATION: n=a,t,g or c
US-09-873-367C-1060

Query Match      98.1%; Score 83.4; DB 3; Length 66109;
Best Local Similarity 98.8%; Pred. No. 1.4e-19;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 37840 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 37899

Qy 61 CTGGGTGCCCCACTAAGTAGAATAA 85
```

Db 37900 CTGGGTGCCCCCACTAAGTAGAATAA 37924

RESULT 14

```

US-10-843-641A-1060
; Sequence 1060, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: Signature Gene Sets
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1060
; LENGTH: 66109
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(66109)
; OTHER INFORMATION: n=a,t,g or c
US-10-843-641A-1060

```

	Query Match	98.1%	Score 83.4	DB 9	Length 66109
	Best Local Similarity	98.8%	Pred. No. 1.4e-19		
	Matches 84	Conservative 0	Mismatches 1	Indels 0	Gaps 0
Qy	1	CGGTGAAGTGCATCTCTCTGGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT	60		
Db	37840	CGGTGAAGTGCATCTCTCTGGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT			
Qy	61	CTGGTGCCCCCACTAAGTAGAAAAA	85		
Db	37900	CTGGTGCCCCCACTAAGTAGAAATAA	37924		

RESULT 15

```

US-09-880-107-3814
; Sequence 3814, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44321-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14

```

```

; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3814
; LENGTH: 198285
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X87344
US-09-880-107-3814

Query Match          98.1%; Score 83.4; DB 3; Length 198285;
Best Local Similarity 98.8%; Pred. No. 2e-19;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  CGGTGAAGTGCATCTTCTGCGTGTCTCTATTTTGAACAAGCATTTCCCCAGGGAAGTTT 60
        |||
Db      113629  CGGTGAAGTGCATCTTCTGCGTGTCTCTATTTTGAACAAGCATTTCCCCAGGGAAGTTT 113688

QY      61  CTGGGTGCCCCCACTAAGTAGAAAAA 85
        |||
Db      113689  CTGGGTGCCCCCACTAAGTAGAATAA 113713

Search completed: December 6, 2005, 22:57:17
Job time : 216.652 secs

```

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 342.427 Seconds
(without alignments)
4099.008 Million cell updates/sec

Title: US-09-980-046B-9
Perfect score: 30
Sequence: 1 cgggaataaaggctgttgaagacaaaaa 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1:	gb_est1:*
2:	gb_est2:*
3:	gb_est3:*
4:	gb_hc:*
5:	gb_est4:*
6:	gb_est5:*
7:	gb_est6:*
8:	gb_est7:*
9:	gb_ga1:*
10:	gb_ga2:*
11:	gb_ga3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	123	6	CD520437 AGENCOURT
2	30	100.0	688	2	BG713896 602673757
3	28.4	94.7	110	3	BM055401 ie94d09.y
4	28.4	94.7	157	2	BI052127 PM0-GN040
5	28.4	94.7	166	1	AI401069 th22a01.x
6	28.4	94.7	203	1	AI342093 q898h01.x
7	28.4	94.7	241	5	BUS35679 AGENCOURT
8	28.4	94.7	245	6	CA426383 UI-H-FEI-
9	28.4	94.7	273	2	BF447271 7p47a11.x
10	28.4	94.7	274	1	AI400897 th30f11.x
11	28.4	94.7	299	3	BQ013504 UI-1-BC1p
12	28.4	94.7	300	3	BQ011929 UI-1-BC1p
13	28.4	94.7	327	1	AA133211 2k97g08.s
14	28.4	94.7	332	6	CF529736 UI-1-BC1p
15	28.4	94.7	349	5	BUS57786 AGENCOURT
16	28.4	94.7	377	5	BUI93594 AGENCOURT
17	28.4	94.7	380	3	BQ028373 UI-H-CO0-
18	28.4	94.7	412	1	AI952448 wx73h02.x
19	28.4	94.7	420	1	AW117898 xe38e06.x
20	28.4	94.7	432	1	AI336331 qt43f06.x
21	28.4	94.7	432	6	CB052146 NISC gl06
22	28.4	94.7	433	1	AI086693 oz64h09.x

23	28.4	94.7	451	1	AJ708390
24	28.4	94.7	457	5	BUI74042
c	28.4	94.7	465	2	BE856455 7f84b05.x
c	28.4	94.7	507	1	AW080368
c	28.4	94.7	542	3	BM999267
28	28.4	94.7	544	5	BUS96408
29	28.4	94.7	556	5	BUS21125
c	28.4	94.7	571	1	AA587336
c	28.4	94.7	575	1	AA149615
c	28.4	94.7	577	1	AA165422
c	28.4	94.7	584	1	AA583402
c	28.4	94.7	592	1	AW001496
c	28.4	94.7	595	1	AA587899
c	28.4	94.7	598	1	AW439384
c	28.4	94.7	605	7	CR992231
c	28.4	94.7	618	5	BUI734305
c	28.4	94.7	625	5	BUI737708
c	28.4	94.7	629	5	BUE21270
c	28.4	94.7	629	7	CR999119
41	28.4	94.7	632	5	BQ946113
42	28.4	94.7	633	5	BUI956376
43	28.4	94.7	637	5	BQ437232
44	28.4	94.7	637	5	BUI84485
45	28.4	94.7	637	5	BUI84485

ALIGNMENTS

RESULT 1
CD520437
LOCUS
DEFINITION AGENCOURT 14354662 NIH MGC 191 Homo sapiens cDNA clone
IMAGE: 30411401 5', mRNA sequence.
CD520437
CD520437.1 GI:31452155
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 123)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
JOURNAL
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Narayan Bhat
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NUCM203 row: b column: 18
High quality sequence stop: 123.
Location/Qualifiers
1. 123
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30411401"
/tissue_type="Pooled"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NIH_MGC_191"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc) ;
Site 2: SfiI (ggcgcctggcc); Library is oligo-dT primed
and directionally cloned. PBMC - Peripheral Blood
Mononuclear Cells. RNA was pooled from 3/6hour stimulation
with PMA adn Ionomycin. 5' and 3' adaptors were used in

CD520437 123 bp mRNA linear EST 06-JUN-2003
AGENCOURT 14354662 NIH MGC 191 Homo sapiens cDNA clone
IMAGE: 30411401 5', mRNA sequence.

CD520437
CD520437.1 GI:31452155
EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

1 (bases 1 to 123)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

JOURNAL

COMMENT

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Narayan Bhat

CDNA Library Preparation: CLONTECH Laboratories, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: NUCM203 row: b column: 18

High quality sequence stop: 123.

Location/Qualifiers

1. 123

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30411401"

/tissue_type="Pooled"

/lab_host="DH10B (TI phage-resistant)"

/clone_lib="NIH_MGC_191"

/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc) ;

Site 2: SfiI (ggcgcctggcc); Library is oligo-dT primed

and directionally cloned. PBMC - Peripheral Blood

Mononuclear Cells. RNA was pooled from 3/6hour stimulation

with PMA adn Ionomycin. 5' and 3' adaptors were used in

cloning as follows: 5' adaptor sequence:
5'-CAGGCCATTATGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGCGCGCATG-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.69
kb (range 0.70-5.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGACAAAA 30
Db 63 CGGAATAAAGGCTGTTGTAAGACAAAA 92

RESULT 2

LOCUS BG7113896 688 bp mRNA linear EST 08-MAY-2001
DEFINITION 602673757F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4796078 5',
mRNA sequence.

ACCESSION BG7113896 GI:13992827

VERSION BG7113896

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 688)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM10680 row: b column: 15

High quality sequence stop: 664.

Location/Qualifiers

1..688

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4796078"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/clone_lib="NIH_MGC_96"

/notes="Organ: Brain; Vector: pBluescriptR (modified

pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI

(gtcgag); Oligo-dT primed using primer

5'-TTTTTTTTTTTTTTVN-3', size-selected for average

insert size 2.3 kb and normalized to ROT 5. This is a

primary library enriched for full-length clones and

constructed using the Cap-trapper method (Carninci, in

preparation). Library constructed by M. Brownstein

(NIH/NHGRI, National Institutes of Health). Note: this is

ORIGIN

Query Match 100.0%; Score 30; DB 2; Length 688;
Best Local Similarity 100.0%; Pred. No. 0.69;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGACAAAA 30
Db 637 CGGAATAAAGGCTGTTGTAAGACAAAA 666

RESULT 3

LOCUS BM055401

DEFINITION

ie94d09.y1 Melton Normalized Human Islet 4 M4-HIS 1 Homo sapiens
cDNA clone IMAGE:5674433 5', mRNA sequence.

ACCESSION BM055401

VERSION BM055401.1 GI:16813271

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 110)

Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,

Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,

Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,

Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,

Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarishvili, R.,

Williams, T., Jackson, Y. and Bowers, Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Other ESTs: ie94d09.x1

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by:

Washington University Genome Sequencing Center For information on

obtaining a clone please contact: Juliana Brown

(brownj@fas.harvard.edu) This sequence now available from the IMAGE

consortium, for clone orders contact: info@image.llnl.gov.

Location/Qualifiers

1..110

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5674433"

/sex="Both"

/tissue_type="Islets of Langerhans"

/dev_stage="Adult"

/lab_host="DH10B"

/clone_lib="Melton Normalized Human Islet 4 M4-HIS 1"

/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;

Site_2: Sal 1; Starting library constructed using

SuperScript Plasmid Library kit (Life Technologies). cDNA

made by oligo-dT priming. Size-selected by column

fractionation; average insert size 1.08 kb. Library was

amplified once on solid support and plasmid DNA from

library was prepared. The library DNA was normalized by

method #4 from Bonaldo, Lennon, and Soares 1996 Genome

Research 6:791-806; 0.5 microgram single-stranded library

plasmid DNA was mixed with 5 micrograms PCR product

representing library inserts and hybridized to an Ecot of

20. Single-stranded (unhybridized) plasmids were isolated

by hydroxyapatite chromatography and used to make this

library."

Query Match 94.7%; Score 28.4; DB 3; Length 110;

Best Local Similarity 96.7%; Pred. No. 2.5;

Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGACAAAA 30

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco.

FEATURES

source

1. .203
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1946161"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP GC4"
/notes="vector: pT73D-Fac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from 3 pooled germ cell tumors, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT773 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 94.7%; Score 28.4; DB 1; Length 203;
Best Local Similarity 96.7%; Pred. No. 2.6;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGCGCTGTTGTAAGACAAAA 30
|||||
Db 52 CGGAATAAAGCGCTGTTGTAAGACAAAA 23
|||||

RESULT 7
BU535679
LOCUS AGENCOURT 10217871 NIH MGC 107 Homo sapiens cDNA clone
DEFINITION IMAGE:6563371 5', mRNA sequence.
ACCESSION BU535679
VERSION BU535679.1 GI:22846120
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 241)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLCM2740 row: 0 column: 19
High quality sequence stop: 240.

FEATURES

source

1. .241
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6563371"

/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_107"
/notes="Organ: breast; Vector: pOTB7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 94.7%; Score 28.4; DB 5; Length 241;
Best Local Similarity 96.7%; Pred. No. 2.7;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGCGCTGTTGTAAGACAAAA 30
|||||
Db 182 CGGAATAAAGCGCTGTTGTAAGACAAAA 211
|||||

RESULT 8
CA426383/c
LOCUS CA426383
DEFINITION UI-H-FEI-beh-g-13-0-UI.81 NCI CGAP_FEI Homo sapiens cDNA clone
ACCESSION CA426383
VERSION CA426383.1 GI:24789109
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 245)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue Procurement: James Martin
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source

1. .245
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FEI-beh-g-13-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FEI"
/notes="Organ: Chondrosarcoma; Vector: pT773-Pac (Pharmacia) with a modified polylinker. Site_1: EcoRI; Site_2: Not I; NCI CGAP_FEI is a normalized cDNA library derived from a pool of mRNA obtained from 3 cell lines from grade II chondrosarcoma tissues. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoRI adaptor, digested with Not I, and cloned directionally into pT773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the

(dnt)18 tail. The sequence tag for this library is CGCTACGGAC. The cell lines were provided by Dr James Martin from the University of Iowa.
TAG_TISSUE=Human grade 2 chondrosarcoma cell line pool
TAG_Lib=UI-H-FE1
TAG_SEQ=CGCTACGGAC

ORIGIN

Query Match 94.7%; Score 28.4; DB 6; Length 245;
Best Local Similarity 96.7%; Pred. No. 2.7; Indels 0; Gaps 0;
Matches 29; Conservative 0; Mismatches 1;

Qy 1 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 30
|||||
Db 41 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 12
|||||

RESULT 9

BF447271/c 273 bp mRNA linear EST 01-DEC-2000
LOCUS 7p47a11.x1 NCI CGAP Pr28 Homo sapiens cDNA clone IMAGE:3648693 3'
DEFINITION similar to gb:UERP_HUMAN P05161 UBIQUITIN CROSS-REACTIVE PROTEIN
PRECUSOR ;, mRNA sequence.

ACCESSION

BF447271 GI:11512409

VERSION

EST.

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

1 (bases 1 to 273)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov

Seq primer: -40UP from Gibco.

Location/Qualifiers

1. .273

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3648693"

/sex="male"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="NCI-CGAP Pr28"

/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI CGAP Pr28 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clonoids
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 94.7%; Score 28.4; DB 2; Length 273;
Best Local Similarity 96.7%; Pred. No. 2.7;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 30

Db 32 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 3
|||||

RESULT 10

AI400897/c

LOCUS

DEFINITION

similar to gb:M13755 INTERFERON-INDUCED 17 KD PROTEIN (HUMAN);,
mRNA sequence.

AI400897 274 bp mRNA linear EST 30-MAR-1999

AI400897.1 GI:4243984

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 274)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 359 Std Error: 0.00

Seq primer: -40UP from Gibco.

Location/Qualifiers

1. .274

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2119821"

/sex="male"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="NCI CGAP Pr28"

/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI CGAP Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clonoids
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 94.7%; Score 28.4; DB 1; Length 274;
Best Local Similarity 96.7%; Pred. No. 2.7; Indels 0; Gaps 0;
Matches 29; Conservative 0; Mismatches 1;

Qy 1 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 30

Db 33 CGGAATAAAGCGCTGTTGTTAAAGACAAAA 4

RESULT 11

BQ013504/c

LOCUS

DEFINITION

UI-1-BC1p-akr-g-03-0-UI.s1 NCI CGAP Pl3 Homo sapiens cDNA clone

UI-1-BC1p-akr-g-03-0-UI 3', mRNA sequence.

ACCESSION BQ013504

VERSION BQ013504.1 GI:19738405

KEYWORDS EST.

BQ013504 299 bp mRNA linear EST 26-MAR-2002
UI-1-BC1p-akr-g-03-0-UI.s1 NCI CGAP Pl3 Homo sapiens cDNA clone
UI-1-BC1p-akr-g-03-0-UI 3', mRNA sequence.

SOURCE
ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 299)
TITLE NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
NATIONAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Dr. Steven Brown
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 Cloning by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES
source Location/Qualifiers
 1..299
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-1-BCIP-akr-g-03-0-UI"
 /tissue_type="Placenta"
 /dev_stage="8-9 weeks"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP Pl3"
 /notes="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Pl3 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are GA, AGGAA. For additional information, contact: Bento Soares, bento-soares@uiowa.edu
 TAG_TISSUE=placenta human 8 week
 TAG_LIB=UI-1-BCIP
 TAG_SEQ=GA"

ORIGIN
 Query Match 94.7%; Score 28.4; DB 3; Length 299;
 Best Local Similarity 96.7%; Pred. No. 2.7;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGGAATAAAGGCTGTTGTAAGACAAAA 30
 |||||||
 DB 40 CGGAATAAAGGCTGTTGTAAGACAAAA 11
 |||||||

RESULT 12
LOCUS BQ011929/c
DEFINITION UI-1-BCIP-ati-b-04-0-UI.s1 NCI CGAP Pl3 Homo sapiens cDNA clone
 UI-1-BCIP-ati-b-04-0-UI 3', mRNA sequence.
ACCESSION BQ011929
VERSION BQ011929.1 GI:19736830
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 300)
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Dr. Steven Brown
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 Cloning by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES
source Location/Qualifiers
 1..300
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-1-BCIP-ati-b-04-0-UI"
 /tissue_type="Placenta"
 /dev_stage="8-9 weeks"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP Pl3"
 /notes="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Pl3 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are GA, AGGAA. For additional information, contact: Bento Soares, bento-soares@uiowa.edu
 TAG_TISSUE=placenta human 8 week
 TAG_LIB=UI-1-BCIP
 TAG_SEQ=GA"

ORIGIN
 Query Match 94.7%; Score 28.4; DB 3; Length 300;
 Best Local Similarity 96.7%; Pred. No. 2.7;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGGAATAAAGGCTGTTGTAAGACAAAA 30
 |||||||
 DB 41 CGGAATAAAGGCTGTTGTAAGACAAAA 12
 |||||||

RESULT 13
LOCUS AAL33211/c
DEFINITION AAL33211/c
 IMAGE:490814.3' similar to gb:W13755 INTERFERON-INDUCED 17 KD PROTEIN (HUMAN);, mRNA sequence.
ACCESSION AAL33211
VERSION AAL33211.1 GI:1689973
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 327)
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chisoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.


```
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_107"
/notes="Organ: breast; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
```

ORIGIN

```
Query Match      94.7%; Score 28.4; DB 5; Length 349;
Best Local Similarity 96.7%; Pred.No. 2.7;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  CGGAATAAAGGCTGTTGTTAAAGACAAAA 30
      |||||||
Db      285 CGGAATAAAGGCTGTTGTTAAAGACAAAA 314
```

Search completed: December 7, 2005, 09:56:03
Job time : 345.527 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 44.9561 Seconds
(without alignments)
4072.612 Million cell updates/sec

Title: US-09-980-046b-8

Perfect score: 103

Sequence: 1 cggacagacggcagcagctcc.....tgagcccttctcacaaaaa 103

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
- 9: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101.4	98.4	427	3	US-09-461-325-107
2	101.4	98.4	427	3	US-10-012-542-107
3	101.4	98.4	427	3	US-10-115-123-107
4	100.4	97.5	1277	3	US-09-461-325-20
5	100.4	97.5	1277	3	US-10-012-542-20
6	100.4	97.5	1277	3	US-10-115-123-20
7	74.4	72.2	280	3	US-09-118-554-58
8	74.4	72.2	280	3	US-09-118-627-58
9	74.4	72.2	280	3	US-09-602-877A-58
10	31.4	30.5	4118	2	US-08-119-125A-3
11	28.6	27.8	601	3	US-09-949-016-87323
12	28.6	27.8	32393	3	US-09-949-016-14203
13	28	27.2	160759	3	US-09-949-016-16514
14	27.6	26.8	1387	3	US-09-663-600A-56
15	27.6	26.8	29636	3	US-09-949-016-12189
16	27.6	26.8	29637	3	US-09-949-016-17173
17	27.6	26.8	29638	3	US-09-949-016-14137
18	27.6	26.8	135667	3	US-09-949-016-15051
19	27.6	26.8	152486	3	US-09-949-016-12869
20	27.4	26.6	601	3	US-09-949-016-147021
21	27.4	26.6	374159	3	US-09-949-016-15868
22	27	26.2	193169	3	US-09-949-016-15091
23	26.8	26.0	113538	3	US-09-949-016-16329
24	26.8	26.0	246230	3	US-09-949-016-17019

ALIGNMENTS

RESULT 1

US-09-461-325-107
Sequence 107, Application US/09461325A
Patent No. 6475753
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 9A Human Secreted Proteins
FILE REFERENCE: P2028P1
CURRENT APPLICATION NUMBER: US/09/461.325A
CURRENT FILING DATE: 1999-12-14
EARLIER APPLICATION NUMBER: PCT/US99/13418
EARLIER FILING DATE: 1999-06-15
EARLIER APPLICATION NUMBER: 60/089,507
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/089,508
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/089,509
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/089,510
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/090,112
EARLIER FILING DATE: 1998-06-22
EARLIER APPLICATION NUMBER: 60/090,113
EARLIER FILING DATE: 1998-06-22
NUMBER OF SEQ ID NOS: 532
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 107
LENGTH: 427
TYPE: DNA
ORGANISM: Homo sapiens
US-09-461-325-107

Query Match 98.4%; Score 101.4; DB 3; Length 427;

Best Local Similarity 99.0%; Pred. No. 1.6e-28;

Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	CGGACAGCGCGAGCTCCAGCTCTGGTTCCCTTCGTTTATCTGTAGATGAA	60
Db	281	CGGACAGCGCGAGCTCCAGCTCTGGTTCCCTTCGTTTATCTGTAGATGAA	340
Qy	61	ATGTTTCCCAATAAAGGGCATGAGCCCTTCCTCAAAAA	103
Db	341	ATGTTTCCCAATAAAGGGCATGAGCCCTTCCTCAAAAA	383

RESULT 2

US-10-012-542-107

Sequence 17020, A
Sequence 17021, A
Sequence 17022, A
Sequence 91, Appl
Sequence 52888, A
Sequence 12578, A
Sequence 16400, A
Sequence 13271, A
Sequence 16, Appl
Sequence 17, Appl
Sequence 40554, A
Sequence 134471, A
Sequence 15473, A
Sequence 17588, A
Sequence 765, App
Sequence 3859, Ap
Sequence 35605, A
Sequence 24, Appl
Sequence 23, Appl
Sequence 25, Appl
Sequence 21, Appl

```
; Sequence 107, Application US/10012542
; Patent No. 6627741
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 107
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-542-107

Query Match          98.4%; Score 101.4; DB 3; Length 427;
Best Local Similarity 99.0%; Pred. No. 1.6e-28;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTCTCGGTTTATTCCTGTTAGAAATGAA 60
Db 281 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTCTCGGTTTATTCCTGTTAGAAATGAA 340

Qy 61 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCTCCACAAAAA 103
Db 341 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCTCCACAAAAA 383

RESULT 4
US-09-461-325-20
; Sequence 20, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/09/461,325A
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418
; EARLIER FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,508
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,509
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,510
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1277
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1207)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1272)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-461-325-20

Query Match          97.5%; Score 100.4; DB 3; Length 1277;
Best Local Similarity 98.1%; Pred. No. 6.2e-28;
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTCTCGGTTTATTCCTGTTAGAAATGAA 60
Db 1130 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTCTCGGTTTATTCCTGTTAGAAATGAA 1189

Qy 61 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCTCCACAAAAA 103
Db 1190 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCTCCACAAAAA 1232

RESULT 5
US-10-012-542-20
; Sequence 20, Application US/10012542
; Patent No. 6627741
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 107
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-542-107

Query Match          98.4%; Score 101.4; DB 3; Length 427;
Best Local Similarity 99.0%; Pred. No. 1.6e-28;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTCTCGGTTTATTCCTGTTAGAAATGAA 60
Db 281 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCCTCTCGGTTTATTCCTGTTAGAAATGAA 340

Qy 61 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCTCCACAAAAA 103
Db 341 ATGGTTCCCATAAATAAGGGGCATGAGCCCTTCTCCACAAAAA 383

RESULT 3
US-10-115-123-107
; Sequence 107, Application US/10115123
; Patent No. 6774216
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029G30AP1D2
; CURRENT APPLICATION NUMBER: US/10/115,123
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: PCT/US99/13418
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 107
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-115-123-107
```



```
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 280
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-118-627-58

Query Match      72.2% Score 74.4; DB 3; Length 280;
Best Local Similarity 98.7%; Pred. No. 2.5e-18;
Matches 75; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 25 TCTGGTTTCCTTCGCGTTTATTCTGTAGAAATGTTCCCATAAATAGGGGCAT 84
Db 266 TCTGGTTTCCTTCGCGTTTATTCTGTAGAAATGTTCCCATAAATAGGGGCAT 207

Qy 85 GAGCCCTTCCTCAAA 100
Db 206 GAGCCCTTCCTCAGA 191

RESULT 9
US-09-602-877A-58/c
; Sequence 58, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602,877A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 280
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-602-877A-58

Query Match      72.2% Score 74.4; DB 3; Length 280;
Best Local Similarity 98.7%; Pred. No. 2.5e-18;
Matches 75; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 25 TCTGGTTTCCTTCGCGTTTATTCTGTAGAAATGTTCCCATAAATAGGGGCAT 84
Db 266 TCTGGTTTCCTTCGCGTTTATTCTGTAGAAATGTTCCCATAAATAGGGGCAT 207

Qy 85 GAGCCCTTCCTCAAA 100
Db 206 GAGCCCTTCCTCAGA 191

RESULT 10
US-08-119-125A-3/c
; Sequence 3, Application US/08119125A
; Patent No. 5610011
; GENERAL INFORMATION:
; APPLICANT: SMITH, Hilda Elizabeth
; APPLICANT: VECHT, Uri
; TITLE OF INVENTION: DNA Sequences which code for Virulence
; TITLE OF INVENTION: Characteristics of Streptococcus suis and parts thereof, poly
; TITLE OF INVENTION: antibodies derived therefrom and the use thereof for the diagn
; TITLE OF INVENTION: protection against infection by S. suis in mammals, including
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centraal Diogeneeskundig Instituut
; STREET: Rdelherweg 15
; CITY: PH Lelystad
; STATE:
; COUNTRY: The Netherlands
```

```
; ZIP: NL-8219
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; OPERATING SYSTEM: MS-DOS v.6.0
; SOFTWARE: WordPerfect v. 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/119,125A
; FILING DATE: 20-SEP-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL92/00054
; FILING DATE: 19-MAR-1992
; APPLICATION NUMBER: NL 9100510
; FILING DATE: 21-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Handal, Anthony H.
; REGISTRATION NUMBER: 26275
; REFERENCE/DOCKET NUMBER: SMITHHE119125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 838-8589
; TELEFAX: (203) 838-8794
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4118 base pairs
; TYPE: Nucleic acid with corresponding amino acids
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus suis type II (pathogenic)
; FEATURE: OTHER INFORMATION: Muramidase released protein (MRP) gene
; FEATURE:
; NAME/KEY: promoter -35 region
; LOCATION: bp 4 to 9
; FEATURE:
; NAME/KEY: promoter -10 region
; LOCATION: bp 29 to 34
; FEATURE:
; NAME/KEY: promoter -35 region
; LOCATION: bp 40 to 45
; FEATURE:
; NAME/KEY: promoter -10 region
; LOCATION: bp 63 to 68
; FEATURE:
; NAME/KEY: ribosome binding site
; LOCATION: bp 147 to 152
; FEATURE:
; NAME/KEY: signal peptide
; LOCATION: bp 159 to 299
; FEATURE:
; NAME/KEY: mature peptide
; LOCATION: bp 300 to 3926
; FEATURE:
; NAME/KEY: proline rich region
; LOCATION: from bp 2757 to 3014
; FEATURE:
; NAME/KEY: repetitive units
; LOCATION: from bp 3015 to 3176, 3423 to 3584 and 3585 to 3743
; FEATURE:
; NAME/KEY: membrane anchor sequence
; LOCATION: from bp 3825 to 3926
; FEATURE:
; NAME/KEY: dyad symmetry regions
; LOCATION: from bp 4069 to 4080 and from bp 4087 to 4098
US-08-119-125A-3

Query Match      30.5% Score 31.4; DB 2; Length 4118;
Best Local Similarity 64.4%; Pred. No. 0.19;
Matches 47; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 29 GTTTCCTTCGCGTTTATTCTGTAGAAATGTTCCCATAAATAGGGGCATGAGC 88
```

```

Db      3778 GTTTCCTCTCAGGTTGTTGGTTTATTAGAAAGTTTCCGATAAAATATGTGTCGTGTTA 3719
      89 CTTCTCCTCACAA 101
      3718 CTTCTTCATCNA 3706

RESULT 11
US-09-949-016-87323
; Sequence 87323, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 87323
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-87323

Query Match          27.8%; Score 28.6; DB 3; Length 601;
Best Local Similarity 67.8%; Pred. No. 0.93; 19; Indels 0; Gaps 0;
Matches 40; Conservative 0; Mismatches 0;

Qy      17 GTCCAGCTCTCGGTTCCCTTCCTCGGTTTATTCTGTAGAAATGAAATGGTTCCCAATAAT 75
      167 GTCACTGATCTGCTTTCTCTCTTCCTTGGCTGTTAGAGGCTCTGATTCAATAACT 225

RESULT 12
US-09-949-016-14203/c
; Sequence 14203, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14203
; LENGTH: 32393
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14203

Query Match          27.8%; Score 28.6; DB 3; Length 32393;
Best Local Similarity 67.8%; Pred. No. 5.6; 19; Indels 0; Gaps 0;
Matches 40; Conservative 0; Mismatches 0;

Qy      17 GTCCAGCTCTCGGTTCCCTTCCTCGGTTTATTCTGTAGAAATGAAATGGTTCCCAATAAT 75

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 462.191 Seconds
(without alignments)
4239.411 Million cell updates/sec

Title: US-09-980-046B-7

Perfect score: 294

Sequence: 1 cggcgatgcggcggttat.....agcgagcacctagacaaaa 294

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_21.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002as.*
- 7: geneseqn2002bs.*
- 8: geneseqn2003as.*
- 9: geneseqn2003bs.*
- 10: geneseqn2003cs.*
- 11: geneseqn2003ds.*
- 12: geneseqn2004as.*
- 13: geneseqn2004bs.*
- 14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	294	100.0	294	4	AAC89738 Human gas
2	269.4	91.6	2369	2	Aav71777 Human mem
3	269.4	91.6	2369	3	Aaz92428 cDNA enco
4	269.4	91.6	2741	13	Adr07477 Full leng
5	269.4	91.6	3047	5	Aas81031 DNA enco
6	267.8	91.1	603	6	Abv94671 Human pan
7	267.8	91.1	3079	12	Adn03645 Antipsori
8	267.8	91.1	3079	13	Actn38103 Tumour-as
9	267.8	91.1	3079	13	Actn04107 Novel hum
10	253.6	86.3	452	9	Ach47801 Human inf
11	244.4	83.1	485	9	Ach29871 Human tes
12	102	34.7	466	13	Adq51104 Novel can
13	35.8	12.2	35973	6	Abk13076 Human amy
14	32.4	11.0	1559	13	Adr31515 Human ade
15	32.4	11.0	2724	6	Abz12539 Arabidops
16	32.2	11.0	2000	8	Ada71938 Rice gene
17	32.2	11.0	120239	13	Abd33244 Murine ca
18	32	10.9	6499	4	Aak86748 Human imm
19	31.6	10.7	105219	11	Acn44286 Human gen

c 20	31.4	10.7	602	12	ADQ19063	Adq19063 Human sof
c 21	31.4	10.7	2000	12	ADJ40696	Adj40696 Plant CDN
c 22	31.4	10.7	3931	5	AAS29195	Aas29195 Genomic s
c 23	31.4	10.7	3931	6	ABS68335	Abs68335 Human DNA
c 24	31.4	10.7	3931	10	ADC25457	Adc25457 Human CDN
c 25	31.4	10.7	50764	11	ACN45090_3	Continuation (4 of
c 26	31.2	10.6	127722	12	ADQ97301	Adq97301 Human can
c 27	31	10.5	5259	13	ADR06908	Adr06908 Full leng
c 28	31	10.5	81656	12	ADQ97876	Adq97876 Human can
c 29	30.8	10.5	567	5	AAS64324	Aas64324 DNA enco
c 30	30.8	10.5	1809	11	ACH99470	Ach99470 Klebsiell
c 31	30.8	10.5	2111	5	ABV27899	Abv27899 Human pro
c 32	30.8	10.5	2111	5	ABV22063	Abv22063 Human pro
c 33	30.8	10.5	3223	4	ABL08588	AbL08588 Drosophil
c 34	30.8	10.5	42999	8	ABX11086	Abx11086 Human rib
c 35	30.6	10.4	238	10	ACD92285	AcD92285 Human col
c 36	30.6	10.4	238	10	ACD92063	AcD92063 Human col
c 37	30.6	10.4	238	10	ACD92284	AcD92284 Human col
c 38	30.6	10.4	238	10	ACD92154	AcD92154 Human col
c 39	30.6	10.4	238	10	ACD92286	AcD92286 Human col
c 40	30.6	10.4	238	10	ACD92228	AcD92228 Human col
c 41	30.6	10.4	238	10	ACD92062	AcD92062 Human col
c 42	30.6	10.4	239	10	ACD92041	AcD92041 Human col
c 43	30.6	10.4	260	12	ADQ05366	Adq05366 Soybean l
c 44	30.6	10.4	277	10	ACD92096	AcD92096 Human col
c 45	30.6	10.4	277	10	ACD92097	AcD92097 Human col

ALIGNMENTS

RESULT 1	
AAC89738	
ID AAC89738 standard; cDNA; 294 BP.	
XX AAC89738;	
XX 12-MAR-2001 (first entry)	
DT Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 7.	
DE Human; cytostatic; immunomodulator; immunostimulant; vulnery;	
KW anti-inflammatory; neuroprotective; antibacterial; gene therapy;	
KW gastrointestinal inflammation; immune system disorder; genetic disorder;	
KW cancer; autoimmune disorder; infection; wound healing; ss.	
XX Homo sapiens.	
OS WO200073324-A2.	
PN 07-DEC-2000.	
PD 01-JUN-2000; 2000WO-US015191.	
PF 01-JUN-1999; 99US-0137059P.	
PR (DIGI-) DIGITAL GENE TECHNOLOGIES INC.	
PA Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Haseel KW;	
PI WPI; 2001-061508/07.	
XX New polynucleotides and polypeptides, useful in gene therapy and in	
PT diagnosing a pathological condition, e.g. for modulating gene expression	
PT in gastrointestinal inflammation, or for treating cancers or genetic	
PT disorders.	
XX Claim 1; Page 85; 108pp; English.	
PS The present sequence is one of a number of isolated human polynucleotides	
XX which are useful in gene therapy, and for diagnosing a pathological	
CC condition or a susceptibility to it. In particular, the polynucleotides	
CC are useful for modulating gene expression in gastrointestinal	

CC inflammation. The polynucleotides are useful for chromosome
 CC identification, controlling gene expression through triple helix
 CC formation or antisense DNA or RNA, or identifying individuals from minute
 CC biological samples using DNA-based identification techniques. The
 CC polynucleotides can also be used as an alternative to restriction
 CC fragment length polymorphism (RFLP), by determining the actual base-by-
 CC base DNA sequences of selected portions of an individual's genome. The
 CC polynucleotides may also be used as molecular weight markers on Southern
 CC gels, as diagnostic probes for the presence of a specific mRNA, as a
 CC probe to substract-out known sequences in the process of discovering novel
 CC polynucleotides, or as an antigen to elicit an immune response. The
 CC polypeptides are useful in diagnostic procedures to detect a disorder.
 CC The polynucleotides and polypeptides are useful for preventing, treating
 CC or ameliorating immune system disorders, genetic disorders, cancers, some
 CC autoimmune disorders, or infections. The polynucleotides and polypeptides
 CC are also useful for differentiating, proliferating or attracting cells,
 CC leading to the regeneration of tissues, especially in wounds or burns.
 CC The polypeptides and polynucleotides may also be used as a food additive
 CC or preservative
 XX
 SQ Sequence 294 BP; 75 A; 87 C; 71 G; 61 T; 0 U; 0 Other;

Query Match 100.0%; Score 294; DB 4; Length 294;
 Best Local Similarity 100.0%; Pred. No. 4.6e-88;
 Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGGCGATGCGGCGGCTTATCCCATGATGCTCTGTCAAGGAGTGACATACCTACT 60
 DB 1 CGGCGATGCGGCGGCTTATCCCATGATGCTCTGTCAAGGAGTGACATACCTACT 60
 QY 61 GAGGAGCCCAACCCCGCTCTGTGATGACCGAATCAGGAGTAGAGTTAACTGCAGCG 120
 DB 61 GAGGAGCCCAACCCCGCTCTGTGATGACCGAATCAGGAGTAGAGTTAACTGCAGCG 120
 QY 121 GAACATGTCAATTTCTCTATTTCTGTGACGCGCCCGCCATGCCCCACCCACCAAGAGG 180
 DB 121 GAACATGTCAATTTCTCTATTTCTGTGACGCGCCCGCCATGCCCCACCCACCAAGAGG 180
 QY 181 CAGGAGGCGCCAGTCATGACCTCGTCTGTGAGGGGTCTGTACAGGTTCTTCTGAAAG 240
 DB 181 CAGGAGGCGCCAGTCATGACCTCGTCTGTGAGGGGTCTGTACAGGTTCTTCTGAAAG 240
 QY 241 CTTTATAGACGTAACAGTTTGCACATGAAATAAAGCGAGCAGCCTAGACAAAAA 294
 DB 241 CTTTATAGACGTAACAGTTTGCACATGAAATAAAGCGAGCAGCCTAGACAAAAA 294

RESULT 2
 ID AAU7177 standard; DNA; 2369 BP.
 XX
 AC AAU7177;
 XX

09-FEB-1999 (first entry)

XX Human membrane fusion protein WDPPro2 encoding DNA.

XX WDPPro1; WDPPro2; WD-40; membrane fusion protein; recombinant; cancer;
 KW beta-transducin; signal transduction; developmental disorder; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 119..1939
 FT /*tag= a
 FT /product= "WDPPro2"

XX US5846777-A.

XX 08-DEC-1998.

XX 26-JUN-1997; 97US-00883534.

PR 26-JUN-1997; 97US-00883534.
 XX (INCY-) INCYTE PHARM INC.

XX Lal P, Bandman O, Corley NC;

XX WPI; 1999-059062/05.

XX P-PSDB; AAW84086.

XX DNA encoding WD-40 proteins - useful for producing recombinant proteins.

XX Claim 10; Fig 3A-F; 42pp; English.

XX This DNA encodes a human membrane fusion protein WDPPro2. The invention
 CC provides polynucleotide sequences (AAU7177 and AAU71777) encoding human
 CC membrane fusion proteins WDPPro1 and WDPPro2. Host cells transformed with
 CC vectors containing the sequences can be used producing recombinant
 CC proteins. WD-40 proteins (collectively designated WDPPro) are homologues
 CC of beta-transducin and function in signal transduction pathways within
 CC the cytoplasm. The WDPPro proteins may be used in the diagnosis,
 CC prevention, and treatment of cancers and developmental disorders
 XX

SQ Sequence 2369 BP; 560 A; 660 C; 695 G; 454 T; 0 U; 0 Other;
 Query Match 91.6%; Score 269.4; DB 2; Length 2369;
 Best Local Similarity 99.6%; Pred. No. 2e-79;
 Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 TCCCATGATGCTCTCTCAAGGAGTGAGCAATCACCTACTGAGGAGCCCCCCCCCGCCT 80
 DB 1898 TCCCATGATGCTCTCTCAAGGAGTGAGCAATCACCTACTGAGGAGCCCCCCCCCGCCT 1957
 QY 81 CTGGATGGACCGAATCAGGAGCTAGAGTTAACTGACGAGCAATGCTCTCTCTATT 140
 DB 1958 CTGGATGGACCGAATCAGGAGCTAGAGTTAACTGACGAGCAATGCTCTCTCTATT 2017
 QY 141 TCTGTGACGCGCCCCCATGCCCCCACCACCAAGAGCAGGAGGCGCCAGTCATGAC 200
 DB 2018 TCTGTGACGCGCCCCCATGCCCCCACCACCAAGAGCAGGAGGCGCCAGTCATGAC 2077
 QY 201 CCTCGCTCTGACGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
 DB 2078 CCTCGCTCTGACGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 2137
 QY 261 CACATGAAAAATAAGCGAGCAGCCTAGACNA 291
 DB 2138 CACATGAAAAATAAGCGAGCAGCCTAGACNA 2168

RESULT 3
 AAU22428

ID AAU22428 standard; cDNA; 2369 BP.

XX AAU22428;

XX 05-JUN-2000 (first entry)

XX cDNA encoding human WD-40 protein, WDPPro2.

XX WDPPro2; human; neutrophil; WD-40 protein; beta-transducin homologue;
 KW signal transduction pathway; cytoplasmic; cancer; developmental disorder;
 KW prevention; treatment; diagnosis; drug screening; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 119..1939
 FT /*tag= a
 FT /product= "Human WDPPro2"

XX US6025464-A.

XX 15-FEB-2000.

```
XX 02-DEC-1998; 98US-00204764.
XX
XX 26-JUN-1997; 97US-00883534.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Lal P, Corley NC, Bandman O;
XX
XX WPI; 2000-181836/16.
XX
XX P-PSDB; AAY77489.
XX
XX New WD-40 protein used for the diagnosis, prevention and treatment of
XX diseases and conditions associated with WD-40 expression e.g. cancer and
XX developmental disorders.
XX
XX Example 5; Fig 3A-F; 42pp; English.
XX
XX The present sequence represents cDNA encoding a human WD-40 protein,
XX WDPro2. cDNAs encoding WD-Pro2 were isolated from a neutrophil cDNA
XX library, this sequence being a consensus. Chemical and structural
XX homology exists between WDPro1 and WD-40 protein (GI 1384131). WD-40
XX proteins are homologues of beta-transducin and function in signal
XX transduction pathways within the cytoplasm. In addition, WDPro2 is
XX expressed in cancers and transformed cells, indicating that it may play a
XX role in cancer and developmental disorders, particularly disorders in
XX which WDPro2 is overexpressed. WDPro2 and WDPro1 (AAV77488) may be used
XX for the prevention, treatment and diagnosis of cancer and developmental
XX disorders e.g., cerebral palsy, or Crohn's disease. The proteins can be
XX used to generate antibodies which can be used to screen libraries of
XX pharmaceutical agents to identify antagonists or agonists of WD-40
XX protein activity
XX
XX Sequence 2369 BP; 560 A; 560 C; 595 G; 454 T; 0 U; 0 Other;
XX
Query Match 91.6%; Score 269.4; DB 3; Length 2369;
Best Local Similarity 99.6%; Pred. No. 2e-79;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 21 TCCCATGATGCTCTGTCAAGAGTGGACATCACTACTGAGGAGGCCACCCCGCCT 80
Db 1898 TCCCATGATGCTCTGTCAAGAGTGGACATCACTACTGAGGAGGCCACCCCGCCT 1957
Qy 81 CTGATGACCGAATCAGGAGTGGAGTTTAACTGCGAGCGGAACATGTCATTCTCTATT 140
Db 1958 CTGATGACCGAATCAGGAGTGGAGTTTAACTGCGAGCGGAACATGTCATTCTCTATT 2017
Qy 141 TCTGTGACGCGCCCGCCATGCCCGCCACCCACCAAGAGGCGAGGCGCCCGATGAC 200
Db 2018 TCTGTGACGCGCCCGCCATGCCCGCCACCCACCAAGAGGCGAGGCGCCCGATGAC 2077
Qy 201 CCTCGTCTCTGACGGGTGCTGTACAGTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db 2078 CCTCGTCTCTGACGGGTGCTGTACAGTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 2137
Qy 261 CACATGAAAATAAAGCGAGCACCTAGACAA 291
Db 2138 CACATGAAAATAAAGCGAGCACCTAGACAA 2168
RESULT 4
ADRO7477
ID ADRO7477 standard; cDNA; 2741 BP.
XX
XX AC ADRO7477;
XX
XX 04-NOV-2004 (first entry)
XX
XX Full length human cDNA useful for treating neurological disease Seq 983.
XX
XX gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;
XX osteoporosis; neurological disease; Alzheimer's disease;
XX Parkinson's disease; dementia; short memory; cancer;
KW sense or motor function; emotional reaction; fear response; panic;
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cyostatic;
KW tranquiliser.
XX
XX Homo sapiens.
XX
XX EP1447413-A2.
XX
XX 18-AUG-2004.
XX
XX 12-FEB-2004; 2004EP-00003145.
XX
XX 14-FEB-2003; 2003JP-00102207.
XX
XX 09-MAY-2003; 2003JP-00131452.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
XX Wakamatsu A, Ishii S, Nagai K, Irie R;
XX
XX WPI; 2004-583265/57.
XX
XX P-PSDB; ADR09433.
XX
XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
XX Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
XX Claim 1; SEQ ID NO 983; 2686pp; English.
XX
XX This invention relates to novel, isolated full length human cDNA
XX molecules and the encoded proteins thereof. Specifically, it refers to
XX cDNA clones obtained by an oligo-capping method, where none of these
XX clones are identical to any known human mRNAs. The present invention
XX describes an immunoassay to identify agonists and antagonists, as well as
XX antibodies, antisense molecules and siRNAs that can all be used to bind
XX to and modulate expression of the cDNA molecules. As such, these
XX molecules are useful for diagnostic markers or therapeutic targets for
XX the various diseases or morbid states. In particular, they are useful in
XX gene therapy for treating osteoporosis, neurological disease, Alzheimer's
XX disease, Parkinson's disease, dementia, short memory and various cancers,
XX as well as for maintaining equilibrium of sense or motor function, and
XX for treating emotional reaction, fear response and panic. Accordingly,
XX they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
XX cyostatic and tranquiliser activities. This polynucleotide is a full
XX length human cDNA sequence of the invention. NOTE: This sequence is not
XX given in the sequence listing of the specification but can be obtained on
XX CD-ROM from the European Patent Office, Vienna Sub-office.
XX
XX Sequence 2741 BP; 643 A; 744 C; 761 G; 593 T; 0 U; 0 Other;
XX
Query Match 91.6%; Score 269.4; DB 13; Length 2741;
Best Local Similarity 99.6%; Pred. No. 2.2e-79;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 21 TCCCATGATGCTCTGTCAAGAGTGGACATCACTACTGAGGAGGCCACCCCGCCT 80
Db 1803 TCCCATGATGCTCTGTCAAGAGTGGACATCACTACTGAGGAGGCCACCCCGCCT 1862
Qy 81 CTGATGACCGAATCAGGAGTGGAGTTTAACTGCGAGCGGAACATGTCATTCTCTATT 140
Db 1863 CTGATGACCGAATCAGGAGTGGAGTTTAACTGCGAGCGGAACATGTCATTCTCTATT 1922
Qy 141 TCTGTGACGCGCCCGCCATGCCCGCCACCCACCAAGAGGCGAGGCGCCCGATGAC 200
Db 1923 TCTGTGACGCGCCCGCCATGCCCGCCACCCACCAAGAGGCGAGGCGCCCGATGAC 1982
Qy 201 CCTCGTCTCTGACGGGTGCTGTACAGTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db 1983 CCTCGTCTCTGACGGGTGCTGTACAGTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 2042
Qy 261 CACATGAAAATAAAGCGAGCACCTAGACAA 291
Db 2043 CACATGAAAATAAAGCGAGCACCTAGACAA 2073
```



```
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 35013; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
XX Sequence 452 BP; 122 A; 131 C; 108 G; 91 T; 0 U; 0 Other;
XX
XX Query Match 86.3%; Score 253.6; DB 9; Length 452;
XX Best Local Similarity 98.2%; Pred. No. 2e-74;
XX Matches 267; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
XX
XX QY 21 TCCCATGATGCTCTGTCAAGAGTGGACAATCACCTACTGAGGAGCCCCACCCCGCCT 80
XX Db TCCCATGATGCTCTGTCAAGAGTGGACAATCACCTACTGAGGAGCCCCACCCCGCCT 188
XX
XX QY 81 CTGGATGACCGAATCAGGACTAGAGTTTAATCGACGGGACATGTCATTCTCTATT 140
XX Db CTGGATGACCGAATCAGGACTAGAGTTTAATCGACGGGACATGTCATTCTCTATT 248
XX
XX QY 141 TCTGTGACGCGCCGCCCATGCCCCACCCACCAAGAGGAGGAGGCCCATGTCATGAC 200
XX Db TCTGTGACGCGCCGCCCATGCCCCACCCACCAAGAGGAGGAGGCCCATGTCATGAC 308
XX
XX QY 201 CCTCGTCTCTGACGGGTGTGTATACAC-GTTTCTCTGAAAGCTTTTAGACAGTAACAGTTT 259
XX Db CCTCGTCTCTGACGGGTGTGTATACACAGTTTCTCTGAAAGCTTTTAGACAGTAACAGTTT 368
XX
XX QY 260 GCACATGAAAAATAAGCGAGCACCTAGACAA 291
XX Db GCACATGAAAAATAAGCGAGCACCTAAACAA 400
XX
XX RESULT 11
XX ACH29871
XX ID ACH29871 standard; cDNA; 485 BP.
XX
XX AC ACH29871;
XX
XX DT 13-OCT-2003 (first entry)
XX
XX DE Human testis cDNA #257.
XX
XX KW genome mapping; biodiversity; genetic disorder.
XX OS Homo sapiens.
XX PN US2003073623-A1.
XX PD 17-APR-2003.
XX PF 30-JUL-2001; 2001US-00918995.
XX PR 30-JUL-2001; 2001US-00918995.
XX PA (DRMA/) DRMANAC R T.
XX PA (LABA/) LABAT I.
XX PA (STAC/) STACHE-CRAIN B.
XX PA (DICK/) DICKSON M C.
XX PA (JONE/) JONES L W.
XX
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 17083; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
XX Sequence 485 BP; 122 A; 147 C; 122 G; 91 T; 0 U; 3 Other;
XX
XX Query Match 83.1%; Score 244.4; DB 9; Length 485;
XX Best Local Similarity 99.6%; Pred. No. 2.6e-71;
XX Matches 245; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 46 GGACAATCACCTACTAGGAGCCCCACCCCGCTCTGGATGGACCCGAATCAGGACTAG 105
XX Db GGACAATCACCTACTAGGAGCCCCACCCCGCTCTGGATGGACCCGAATCAGGACTAG 134
XX
XX QY 106 AGTTTAATCGACGGGACATGTCATTCTCTATTCTTCTGTCAGCGGCCCATGCCCA 165
XX Db AGTTTAATCGACGGGACATGTCATTCTCTATTCTTCTGTCAGCGGCCCATGCCCA 194
XX
XX QY 166 CCCACCAACAGAGGAGGAGGCCCATGTCATGACCCCTCGTCTCTGACGGGTGTCTGTAC 225
XX Db CCCACCAACAGAGGAGGAGGCCCATGTCATGACCCCTCGTCTCTGACGGGTGTCTGTAC 254
XX
XX QY 226 ACGTTCCTCTGAAAGCTTTTAGACAGTAACAGTTTGGACATGAAAAATAAGCGGACCT 285
XX Db ACGTTCTCTGAAAGCTTTTAGACAGTAACAGTTTGGACATGAAAAATAAGCGGACCT 314
XX
XX QY 286 AGACAA 291
XX Db 315 AAACAA 320
```

RESULT 12

ADQ51104
ID ADQ51104 standard; DNA; 466 BP.

AC ADQ51104;

DT 21-OCT-2004 (first entry)

XX Novel canine microarray-related DNA sequence SeqID2406.

XX canine microarray; drug screening; toxicity assay;
XX environmental pollutant; cellular response; gene expression profile;
XX toxic response; liver necrosis; fatty liver disease;
XX protein adduct formation; hepatitis; dog; ds.

XX Canis familiaris.

XX WO2004063324-A2.

XX 29-JUL-2004.

XX 05-MAY-2003; 2003WO-US013853.

XX 03-MAY-2002; 2002US-0377240P.

XX (GENE-) GENE LOGIC INC.

XX (PFIZ) PFIZER PROD INC.

XX Diggins JC, Porter M, Wei T;

XX WPI; 2004-561890/54.

XX New isolated nucleic acid molecule, useful for drug screening and
XX toxicity assays or for assessing the impact, including toxicity, of a
XX compound, pharmaceutical agent or environmental pollutant on a cell or
XX living organism.

XX Claim 1; SEQ ID NO 2406; 41pp; English.

XX This invention is related to a novel isolated canine nucleic acid
XX sequences and the construction of canine microarrays containing a
XX significant portion of the canine genome. The isolated canine nucleic
XX acid sequences of the invention may be useful for drug screening and
XX toxicity assays. The invention is therefore useful for assessing the
XX impact, including toxicity, of a compound, pharmaceutical agent or
XX environmental pollutant on a cell or living organism. The methods are
XX useful for detecting genes that are up- or down-regulated in canines in a
XX disease state. The sequences are useful as diagnostic agents or markers
XX to detect a cellular response in a sample individually or as part of a
XX gene expression profile. It is also useful as a target for agents that
XX modulate gene expression or activity. The database is useful for
XX producing electronic Northern blots that allow the user to determine the cell
XX type or tissue in which a given gene is expressed and to allow
XX determination of the abundance or expression level of a given gene in a
XX particular tissue or cell. The methods are useful for determining the
XX similarity of a toxic response to one or more individual compounds. The
XX methods are useful for predicting at least one toxic response or the
XX likelihood that a compound or test agent will induce various specific
XX pathologies such as those of the liver (liver necrosis, fatty liver
XX disease, protein adduct formation or hepatitis), those of the kidney,
XX heart, brain or testes, or other pathologies associated with at least one
XX of the toxins. The methods are also useful for predicting or elucidating
XX the potential cellular pathways influenced, induced or modulated by the
XX compound or test agent due to the similarity of the expression profile
XX compared to the profile induced by a known toxin. The present sequence is
XX that of a canine DNA sequence which was claimed for use during the
XX production of a canine microarray of the invention.

XX Sequence 466 BP; 102 A; 143 C; 113 G; 100 T; 0 U; 8 Other;

Query Match

Best Local Similarity 34.7%; Score 102; DB 13; Length 466;

Pred. No. 1.6e-23;

	Matches	182;	Conservative	0;	Mismatches	44;	Indels	16;	Gaps	4;
QY	21	TCCCATGATGCTCTCTCAAGGAGTGGACATACCTACTGAGAGCCCCACCCCGCCT	80							
DB	230	TCCCATGATGCTCTCTCAAGGAGTGGACATACCTACTGAGAGCCCCACCCCGCCT	284							
QY	81	CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTATTCTC	136							
DB	285	TTGGACAGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTATTCTC	344							
QY	137	TATTCTGTGACGCGCCCGCCATGCGCCCGCCACCCACCAAGAGCGAGGCGCCAGTCA	196							
DB	345	TATTCTGTGATGCGCGCCCGCCATGCGCCCGCCACCCACCAAGAGCGAGGCGCCAGTCA	404							
QY	197	TGACC-CTGCTCTCTGAGGGTG-----TCTGTACAGTTCTTCTGAAAGCTTTAGACA	249							
DB	405	TGCCCTCATCTCTCTGAGGGTGTTCTCTATCTGTACACGCTCTTGTGAAGCTTTAGACA	464							
QY	250	GT 251								
DB	465	GT 466								

RESULT 13

ABK13076/c

ID ABK13076 standard; DNA; 35973 BP.

XX AC ABK13076;

XX 23-APR-2002 (first entry)

XX Human amyloid beta precursor protein binding protein 1 gene.

XX Human; amyloid beta precursor protein binding protein 1; APPBP1; gene;
XX Alzheimer's disease; transgenic animal; platelet aggregation;
XX single nucleotide polymorphism; SNP; ds.

XX Homo sapiens.

	Key	Location/Qualifiers
FT	variation	replace(3436,G)
FT		/tag= a
FT		/standard name= "Single nucleotide polymorphism"
FT	variation	replace(3474,C)
FT		/tag= b
FT		/standard name= "Single nucleotide polymorphism"
FT	CDS	4085..31975
FT		/tag= c
FT		/product= "APPBP1"
FT		/note= "Amyloid beta precursor protein binding protein 1"
FT	variation	replace(11314,G)
FT		/tag= d
FT		/standard name= "Single nucleotide polymorphism"
FT	variation	replace(18123,G)
FT		/tag= e
FT		/standard name= "Single nucleotide polymorphism"
FT	variation	replace(18571,A)
FT		/tag= f
FT		/standard name= "Single nucleotide polymorphism"
FT	variation	replace(21229,A)
FT		/tag= g
FT		/standard name= "Single nucleotide polymorphism"
FT	variation	replace(21312,A)
FT		/tag= h
FT		/standard name= "Single nucleotide polymorphism"
FT	variation	replace(24300,A)
FT		/tag= i
FT		/standard name= "Single nucleotide polymorphism"
FT	variation	replace(26093,A)
FT		/tag= j
FT		/standard name= "Single nucleotide polymorphism"
FT	variation	replace(29334,C)
FT		/tag= k


```
FT /standard_name= "Single nucleotide polymorphism"
FT replace(31499,T)
FT /*tag= 1
FT /standard_name= "Single nucleotide polymorphism"
FT replace(31582,G)
FT /*tag= m
FT /standard_name= "Single nucleotide polymorphism"
PN WO200202820-A1.
PN
XX
PD 10-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-US020951.
XX
XX 30-JUN-2000; 2000US-0215511P.
XX
XX (GENA-) GENAISSANCE PHARM INC.
XX
XX Anastasio AE, Chew A, Choi JY, Kazemi A, Koshy B, Sausker EA;
XX Stephens CJ;
XX
XX WPI; 2002-164539/21.
XX P-PSDB; AAU11030.
XX
XX Amyloid beta precursor protein binding protein 159 kD (APPBP1) gene
XX polymorphic variants, useful e.g. in studying the expression and function
XX of APPBP1 and screening candidate drugs for treating Alzheimer's disease.
XX
XX Claim 1; Fig 1; 104pp; English.
XX
XX The invention relates to an isolated polypeptide comprising a sequence
XX which is a polymorphic variant of a reference sequence for the amyloid
XX beta precursor protein binding protein 1, 59kD (APPBP1) protein or its
XX fragment. The polymorphic variants are useful in studying the expression
XX and function of APPBP1, in expressing APPBP1 protein for use in screening
XX for candidate drugs to treat diseases related to APPBP1 activity, in
XX studying the effect of the variation on the biological activity of
XX APPBP1, and the binding affinity of candidate drugs targeting APPBP1 for
XX the treatment of disorders such as Alzheimer's disease. The haplotyping
XX methods are useful in validating APPBP1 as a candidate target for
XX treating a specific condition or disease predicted to be associated with
XX APPBP1 activity, or in the design of clinical trials of candidate drugs
XX for treating a specific condition or disease associated with APPBP1
XX activity. The transgenic animals are useful for studying expression of
XX the APPBP1 isogenes in vivo, for in vivo screening and testing of drugs
XX targeted against APPBP1 protein, and for testing the efficacy of
XX therapeutic agents and compounds for disorders related to platelet
XX aggregation in a biological system. The present sequence represents the
XX human APPBP1 gene as described in the method of the invention
XX
XX Sequence 35973 BP; 9909 A; 7327 C; 7959 G; 10766 T; 0 U; 12 Other;
XX
XX Query Match 12.2%; Score 35.8; DB 6; Length 35973;
XX Best Local Similarity 52.3%; Pred. No. 1.8;
XX Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
XX
XX 65 AGCCCCCCCCCTCTGGTGGACCAATCAGGACTAGAGTTTAACTCAGCGGAAC 124
XX |||||
XX 7063 AGCAATTTCTCGCCCTCTAGAGTTGCTGGGACTACAGGCACAGCACTGTACC 7004
XX |||||
XX 125 ATGTCATTTCTATTCTGTGACGCGCCCGCCATGCCCGCCCGCCACCAAGAGGCGAG 184
XX |||||
XX 7003 TGGCTTCATTTGTTTGTGCTGAGCATCATCTTTCCACTGATGCTCCATA 6944
XX |||||
XX 185 AGGGCCCAAGTCATGACCTCGCTCTCTGCAGG 215
XX |||||
XX 6943 AGGGCCCAAGACAGCTGTCAAATGTATGAGG 6913
XX |||||
XX
XX RESULT 14
XX ADR31515
XX ID ADR31515 standard; cDNA; 1559 BP.
XX
```

```
AC ADR31515;
XX
XX 02-DEC-2004 (first entry)
XX
XX Human adenosine deaminase cDNA.
XX
XX Human; adenosine deaminase; ADA; dideoxyinosine; ddi; AIDS; AZT;
XX 3'-azido-2',3'-dideoxythymidine; dideoxyadenosine deaminase; dda; ss;
XX gene.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 122..1213
XX FT /*tag= a
XX FT /product= "Adenosine deaminase"
XX
XX US2004175804-A1.
XX
XX 09-SEP-2004.
XX
XX 26-FEB-2004; 2004US-00787284.
XX
XX 04-MAR-2003; 2003US-0451842P.
XX
XX (SKON/) SKONEZNY P M.
XX (POLI/) POLITINO M.
XX (LIUS/) LIU S W
XX (BOYL/) BOYLE A W.
XX (CHEN/) CHEN J G.
XX (STEI/) STEIN G L.
XX (FRAN/) FRANCESCHINI T.
XX (ANDE/) ANDERSON W L.
XX
XX Skonezny PM, Politino M, Liu SW, Boyle AW, Chen JG, Stein GL;
XX Franceschini T, Anderson WL;
XX WPI; 2004-652027/63.
XX P-PSDB; ADR31514.
XX
XX Making dideoxyinosine (ddi), by obtaining enzyme expressing
XX dideoxyadenosine (dda) deaminase activity, immobilizing enzyme onto
XX insoluble support, contacting enzyme with dda solution to produce ddi
XX solution, isolating ddi.
XX
XX Claim 9; SEQ ID NO 2; 13pp; English.
XX
XX The invention relates to making 2',3'-dideoxyinosine (ddi), involves
XX obtaining an enzyme expressing dideoxyadenosine (dda) deaminase activity,
XX immobilizing the enzyme onto an insoluble support, contacting the enzyme
XX with a dda solution of at least 1% weight volume dda in water for a time
XX and under conditions to produce a ddi solution, and isolating the ddi
XX from the ddi solution. The enzyme is human adenosine deaminase (ADA)
XX appearing as ADR31514 and is encoded by the cDNA appearing as ADR31515 or
XX the E. coli-codon optimised cDNA appearing as ADR31516. The method is
XX useful for making 2',3'-dideoxyinosine (ddi) and provides a yield of at
XX least about 96% ddi that is at least about 99% pure. The ddi mother
XX liquor can be reused in subsequent runs. The human ADA is more stable
XX over a product pH range. ddi is used with AZT (3'-azido-2',3'-
XX dideoxythymidine) to treat AIDS. The present sequence is the wild-type
XX cDNA encoding human ADA.
XX
XX Sequence 1559 BP; 399 A; 448 C; 433 G; 279 T; 0 U; 0 Other;
XX
XX Query Match 11.0%; Score 32.4; DB 13; Length 1559;
XX Best Local Similarity 47.1%; Pred. No. 6.4;
XX Matches 99; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
XX
XX 85 ATGGACCGAATCAGGACTAGAGTTTAACTCAGCGGAACATGTCATTTCTATTCTG 144
XX |||||
XX 1344 ATGTGTCATTTCTGCACACACGATATACCTCGGATGCGCGGTCATCTCTGATGATG 1403
XX |||||
XX 145 TGACGCGCGCCCGCCCATGCGCCCGCCACCCACCAAGAGCGAGGGGCCCATGACCCCTC 204
XX |||||
```

```
Db      1404  TGCCCTGGCCAGGAGCCAGCGCCCTTCGCACATGGGCATGGTTGAATCTGAACCCCTCCTT 1463
Qy      205  GTCTCTCAGGGTGTCTGTACACGTTCTTCTGAAGCTTTAGACAGTAACAGTTTGACACA 264
Db      1464  CTGTGGCAACTGTGTACTGAAATCTGTGTCTCAATAAAGAGCCCATGGCTGGTGGCATG 1523
Qy      265  TGAATAATAAAGCGAGCACCTAGACAAAAA 294
Db      1524  CAAAAAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1553
```

RESULT 15
ABZ12539/c
ID ABZ12539 standard; DNA; 2724 BP.

XX ABZ12539;

XX 21-JAN-2003 (first entry)

XX Arabidopsis thaliana stress regulated gene SEQ ID NO 344.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX Arabidopsis thaliana.

XX WO200216655-A2.

XX 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026685.

XX 24-AUG-2000; 2000US-0227866P.

XX 26-JAN-2001; 2001US-0264647P.

XX 22-JUN-2001; 2001US-0300111P.

XX (SCRI) SCRIPPS RES INST.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

XX Identifying a stress condition to which a plant cell has been exposed and

XX producing plants with increased tolerance to these abiotic stresses.

XX Claim 144; SEQ ID NO 344; 577pp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant
XX cell has been exposed, comprising: (a) contacting nucleic acid
XX representative of expressed polynucleotides in the plant cell with an
XX array or probes representative of the plant cell genome; and (b)
XX detecting a profile of expressed polynucleotides in the plant cell
XX characteristic of a stress response. The method is useful in the
XX production of transgenic plants, cells and seeds and in producing plants
XX with increased tolerance to abiotic stress. The present sequence is that
XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
XX in methods of the invention. Note: The sequence data for this patent is
XX not represented in the printed specification but is based on sequence
XX information supplied to Derwent by the European Patent Office

XX Sequence 2724 BP; 751 A; 558 C; 676 G; 739 T; 0 U; 0 Other;

Query Match 11.0%; Score 32.4; DB 6; Length 2724;
Best Local Similarity 60.0%; Pred. No. 8.1;
Matches 54; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 192 AGTCATGACCTCGCTCTCGAGGCTGTGTACACGTTCTTCTGAAGCTTTAGACAGT 251

Db 2487 AGGAATGATCCCAACACGAGGCTGCTCGATGGATTCTCTCAAAGCTCTTGGCAAT 2428

Qy 252 AACAGTTTGCACATGAAAAATAAAGCGAGC 281

Db 2427 TACAGCTTTTCACTCCCAAAAAAGCAAGGGACC 2398

Search completed: December 7, 2005, 00:03:10
Job time : 465.191 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 140.979 Seconds
(without alignments)
4072.612 Million cell updates/sec

Title: US-09-980-046B-6
Perfect score: 323
Sequence: 1 cggggccaggctatgcct.....gtgcacgtgttgtaaaaa 323

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/6 COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
7: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	311.8	96.5	764	3	US-08-991-789A-283
C 2	311.8	96.5	764	3	US-09-062-451-283
C 3	311.8	96.5	764	3	US-09-289-198-283
C 4	311.8	96.5	764	3	US-09-429-755-283
C 5	311.8	96.5	764	3	US-09-699-295-283
C 6	311.8	96.5	764	3	US-09-534-825A-283
C 7	85	26.3	642	3	US-09-370-838-119
C 8	85	26.3	642	3	US-09-854-133-119
C 9	84.8	26.3	808	3	US-09-023-655-1279
C 10	84.8	26.3	808	3	US-09-461-912A-25
C 11	76	23.5	362	3	US-09-461-912A-2
C 12	52.6	16.3	362	3	US-09-461-912A-1
C 13	50	15.5	50	3	US-10-131-827-2535
C 14	34.6	10.7	289	3	US-09-007-005-17
C 15	34.6	10.7	289	3	US-09-244-796-17
C 16	34.6	10.7	399	3	US-09-621-976-8976
C 17	34.6	10.7	341	3	US-09-949-016-3807
C 18	34.2	10.6	23155	3	US-09-949-016-12366
C 19	34	10.5	25733	3	US-09-902-540-1215
C 20	32.6	10.1	152481	3	US-09-949-016-12521
C 21	32.6	10.1	152798	3	US-09-949-016-12775
C 22	32.6	10.1	152822	3	US-09-949-016-17518
C 23	32.6	10.1	152822	3	US-09-949-016-17519
C 24	32.4	10.0	3306	3	US-09-081-385-10

25	32.4	10.0	3306	3	US-09-752-639-10	Sequence 10, Appl
26	32.4	10.0	3306	3	US-09-712-813-10	Sequence 10, Appl
27	32.4	10.0	3306	3	US-09-700-354A-10	Sequence 10, Appl
C 28	31.6	9.8	9279	3	US-09-487-558B-171	Sequence 171, App
C 29	31.2	9.7	847	3	US-09-533-559-2149	Sequence 2149, App
30	31	9.6	1041	3	US-09-248-796A-1489	Sequence 1489, App
31	31	9.6	1230	3	US-09-252-991A-947	Sequence 947, App
32	31	9.6	1521	3	US-09-252-991A-989	Sequence 989, App
C 33	31	9.6	2201	3	US-09-149-476-245	Sequence 245, App
C 34	30.8	9.5	660	3	US-09-902-540-3995	Sequence 3995, App
35	30.8	9.5	1023	3	US-09-543-681A-1606	Sequence 1606, App
36	30.8	9.5	7218	2	US-08-232-463-14	Sequence 14, Appl
37	30.8	9.5	19862	3	US-09-902-540-1198	Sequence 1198, App
C 38	30.6	9.5	534	3	US-09-593-995-6	Sequence 6, Appl
C 39	30.6	9.5	534	3	US-10-139-667-6	Sequence 6, Appl
40	30.6	9.5	1068	3	US-09-170-496D-129	Sequence 129, App
41	30.6	9.5	1068	3	US-09-170-496D-231	Sequence 231, App
42	30.6	9.5	1161	2	US-08-153-848-31	Sequence 31, Appl
43	30.6	9.5	1161	6	PCT-US93-11153-31	Sequence 31, Appl
44	30.6	9.5	1320	2	US-08-599-252-84	Sequence 84, Appl
45	30.6	9.5	1320	2	US-08-436-074-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1
US-08-991-789A-283/c
; Sequence 283, Application US/08991789A
; Patent No/ 6225054

GENERAL INFORMATION:
APPLICANT: Prudakis, Tony N.
Reed, Steven G.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND DIAGNOSIS OF BREAST CANCER,

NUMBER OF SEQUENCES: 292

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed IP Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/991,789A

FILING DATE: 11-Dec-1997

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Potter, Jane E. R.

REGISTRATION NUMBER: 33,332

REFERENCE/DOCKET NUMBER: 210121.419C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 283:

SEQUENCE CHARACTERISTICS:

LENGTH: 764 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 283:

US-08-991-789A-283

Query Match 96.5%; Score 311.8; DB 3; Length 764;

Best Local Similarity 94.7%; Pred. No. 2.6e-93;

Matches 306; Conservative 14; Mismatches 3; Indels 0; Gaps 0;

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180
Db 210 TATTATGTTACAGATAATACAGGAAAAACGGGTTABTAGTAGCGGCCCATAGCCTGCAA 151
Qy 181 CTTTGCACCTCAGTGTCAATGCTGGCCCTGCAGCTGGGGCTGTGGCCCTGCCCT 240
Db 150 CTTTGCACCTCAGTGTCAATGCTGGCCCTGCAGCTGGGGCTGTGGCCCTGCCCT 91
Qy 241 TGGTCTCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 90 TGGTCTCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 31
Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323
Db 30 AAAGTGCACGTGCTTGTGAAAAA 8

RESULT 4

US-09-429-755-283/c
; Sequence 283, Application US/09429755A
; Patent No. 6656480

GENERAL INFORMATION:

; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; APPLICANT: Mishner, Lynda
; APPLICANT: Retter, Marc W.
; APPLICANT: Dillon, Davin C.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; FILE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER

; FILE REFERENCE: 210121.419C6

; CURRENT APPLICATION NUMBER: US/09/429,755A

; CURRENT FILING DATE: 1999-10-28

; NUMBER OF SEQ ID NOS: 315

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 283

; LENGTH: 764

; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(764)

; OTHER INFORMATION: n = A,T,C or G

US-09-429-755-283

Query Match 96.5%; Score 311.8; DB 3; Length 764;

Best Local Similarity 94.7%; Pred. No. 2.6e-93;

Matches 306; Conservative 14; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTTCTGGG 60
Db 330 CGGGGCMAGGCTATGCTCCACCGCAAGTGCCTGAASATSTGGGCCCTGATTTSTGG 271
Qy 61 CATCTCATGACCAATGGAATCACCTGTACTGTGTTACTGGCTCTGTGACAGTCTACCA 120
Db 270 CATCTCTGACCTTGGATTTCATCTGTDAYTGGWATTGGCTCTGTGACAGTCTACCA 211
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180
Db 210 TATTATGTTACAGATAATACAGGAAAAACGGGTTABTAGTAGCGGCCCATAGCCTGCAA 151
Qy 181 CTTTGCACCTCAGTGTCAATGCTGGCCCTGCAGCTGGGGCTGTGGCCCTGCCCT 240
Db 150 CTTTGCACCTCAGTGTCAATGCTGGCCCTGCAGCTGGGGCTGTGGCCCTGCCCT 91
Qy 241 TGGTCTCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 90 TGGTCTCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 31
Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323
Db 30 AAAGTGCACGTGCTTGTGAAAAA 8

RESULT 5

US-09-699-295-283/c
; Sequence 283, Application US/09699295

GENERAL INFORMATION:

; APPLICANT: Frudakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Mishner, Lynda
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Harlocker, Susan L.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; FILE OF INVENTION: THERAPY AND DIAGNOSIS OF BREAST CANCER

; FILE REFERENCE: 210121.419C10

; CURRENT APPLICATION NUMBER: US/09/699,295

; CURRENT FILING DATE: 2000-10-26

; NUMBER OF SEQ ID NOS: 326

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 283

; LENGTH: 764

; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(764)

; OTHER INFORMATION: n = A,T,C or G

US-09-699-295-283

Query Match 96.5%; Score 311.8; DB 3; Length 764;

Best Local Similarity 94.7%; Pred. No. 2.6e-93;

Matches 306; Conservative 14; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTTCTGGG 60
Db 330 CGGGGCMAGGCTATGCTCCACCGCAAGTGCCTGAASATSTGGGCCCTGATTTSTGG 271
Qy 61 CATCTCATGACCAATGGAATCACCTGTACTGTGTTACTGGCTCTGTGACAGTCTACCA 120
Db 270 CATCTCTGACCTTGGATTTCATCTGTDAYTGGWATTGGCTCTGTGACAGTCTACCA 211
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180
Db 210 TATTATGTTACAGATAATACAGGAAAAACGGGTTABTAGTAGCGGCCCATAGCCTGCAA 151
Qy 181 CTTTGCACCTCAGTGTCAATGCTGGCCCTGCAGCTGGGGCTGTGGCCCTGCCCT 240
Db 150 CTTTGCACCTCAGTGTCAATGCTGGCCCTGCAGCTGGGGCTGTGGCCCTGCCCT 91
Qy 241 TGGTCTCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 90 TGGTCTCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 31
Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323
Db 30 AAAGTGCACGTGCTTGTGAAAAA 8

RESULT 6

US-09-534-825A-283/c

; Sequence 283, Application US/09534825A

GENERAL INFORMATION:

; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; APPLICANT: Mishner, Lynda
; APPLICANT: Retter, Marc W.
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

```
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C7
; CURRENT APPLICATION NUMBER: US/09/534,825A
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 317
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 283
; LENGTH: 764
; TYPE: DNA
; ORGANISM: Homo sapien
; NAME/KEY: misc feature
; LOCATION: (1)-(764)
; OTHER INFORMATION: n = A,T,C or G
US-09-534-825A-283

Query Match 96.5%; Score 311.8; DB 3; Length 764;
Best Local Similarity 94.7%; Pred. No. 2.6e-93;
Matches 306; Conservative 14; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTTCTGGG 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 330 CGGGGCMAGGCSATGCCGCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTTSTGG 271
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 61 CATCTCATGACCAATTGGATTACCCCTGTTACTGCTGTTTGGTATTCGGCTCTGTGACAGTCTACCA 120
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 270 CATCTCATGACCAATTGGATTACCCCTGTTACTGCTGTTTGGTATTCGGCTCTGTGACAGTCTACCA 211
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTACGCCGCCATAGCCTGCAA 180
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 210 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTACGCCGCCATAGCCTGCAA 151
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 181 CCTTTGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 150 CCTTTGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 91
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 241 TGGTCTGCGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 90 TGGTCTGCGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 31
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 30 AAAGTGCACGTGCTTGTGAAAAA 8
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 7
US-09-370-838-119
; Sequence 119, Application US/09370838
; Patent No. 644425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 119
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-370-838-119

Query Match 26.3%; Score 85; DB 3; Length 642;
Best Local Similarity 60.7%; Pred. No. 5.7e-18;
Matches 196; Conservative 0; Mismatches 90; Indels 37; Gaps 2;

Qy 1 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTTCTGGG 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 61 CATCTCATGACCAATTGGATTACCCCTGTTACTGCTGTTTGGTATTCGGCTCTGTGACAGTCTACCA 120
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 392 CATCTCATGACCAATTGCTCTCATGCTATCCAGTGTCTGATCTTCCAGGCTATGGATA 451
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTACGCCGCCATAGCCTGCAA 180
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 452 GATCAGG-----AGGCATCACTAGAGCCAGGAGCTGTGCCCATGACCTGTAT 498
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 181 CCTTTGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 499 CCCAGCTACTCCAACTTCCATTCCTGCGCCCTG-----CCCC 534
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 241 TGGTCTGCGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 535 CGGAGCGGAGTCTGTATCAGCCCTTTATCTCACACGCTTTTCTACAATGGCATTCAAT 594
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 8
US-09-854-133-119
; Sequence 119, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 119
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-854-133-119

Query Match 26.3%; Score 85; DB 3; Length 642;
Best Local Similarity 60.7%; Pred. No. 5.7e-18;
Matches 196; Conservative 0; Mismatches 90; Indels 37; Gaps 2;

Qy 1 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTTCTGGG 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 332 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTTCTGGG 391
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 61 CATCTCATGACCAATTGGATTACCCCTGTTACTGCTGTTTGGTATTCGGCTCTGTGACAGTCTACCA 120
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 392 CATCTCATGACCAATTGCTCTCATGCTATCCAGTGTCTGATCTTCCAGGCTATGGATA 451
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTACGCCGCCATAGCCTGCAA 180
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 452 GATCAGG-----AGGCATCACTAGAGCCAGGAGCTGTGCCCATGACCTGTAT 498
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 181 CCTTTGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 499 CCCAGCTACTCCAACTTCCATTCCTGCGCCCTG-----CCCC 534
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 241 TGGTCTGCGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 535 CGGAGCGGAGTCTGTATCAGCCCTTTATCTCACACGCTTTTCTACAATGGCATTCAAT 594
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db      595 AAGTGCACGTGTTCTGGTGA 617
|||||
RESULT 9
US-09-023-655-1279
; Sequence 1279, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Sellhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1279:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 808 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: G311374
US-09-023-655-1279

Query Match      26.3%; Score 84.8; DB 3; Length 808;
Best Local Similarity 61.0%; Pred. No. 7.4e-18;
Matches 194; Conservative 0; Mismatches 87; Indels 37; Gaps 2;

Qy      1 CGGGGCCAGGCTATGCTCCACGGCAAGTGCCTGAAACATCTGGGCCCTGATTCGGG 60
Db      519 CGGGGCCAGGCTATGCTCCACGGCAAGTGCCTGAAACATCTGGGCCCTGATTCGGG 578

Qy      61 CATCTCATGACCATTTGGATTACCCCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 120
Db      579 CATCTCATGACCATTTGCTGCTCATGCTCATCCAGTGTGATCTTCAGGCTATGGATA 638

Qy      121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180
Db      639 GATCAGG-----AGGCATCACTGAGGCCAGGAGCTCTGGCCCATGACCTGTAT 685

Qy      181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGACAGCTTGGGGCTGTGGCCCTGCCCCCT 240
Db      686 CCCAGCTACTCCAACTTCCCATTTCTCGCCCTG-----CCCC 721

Qy      241 TGGTCTCGCCCTAGATACAGACAGTGTATATCCACACACACCTGTCTACAGTGTCTTCAAT 300
Db      722 CGGAGCGAGTCTGTATACGCCCTTTATCTCACAGCTTTTCTACAAATGGCATTCAAT 781

Qy      301 AAAGTGCACGTGTTGTG 318
Db      782 AAAGTGCACGTGTTTCTG 799

RESULT 10
US-09-461-912A-25
; Sequence 25, Application US/09461912A
; Patent No. 6709855
; GENERAL INFORMATION:
; APPLICANT: Stanton, Lawrence A.
; APPLICANT: White, R. Tyler
; APPLICANT: Damm, Deborah L.
; APPLICANT: Lewicki, John A.
; TITLE OF INVENTION: Methods for detection and use of
; TITLE OF INVENTION: differentially expressed genes in disease states
; FILE REFERENCE: SCIOS.011A
; CURRENT APPLICATION NUMBER: US/09/461,912A
; CURRENT FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,008
; PRIOR FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 808
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (238)....(639)
US-09-461-912A-25

Query Match      26.3%; Score 84.8; DB 3; Length 808;
Best Local Similarity 61.0%; Pred. No. 7.4e-18;
Matches 194; Conservative 0; Mismatches 87; Indels 37; Gaps 2;

Qy      1 CGGGGCCAGGCTATGCTCCACGGCAAGTGCCTGAAACATCTGGGCCCTGATTCGGG 60
Db      519 CGGGGCCAGGCTATGCTCCACGGCAAGTGCCTGAAACATCTGGGCCCTGATTCGGG 578

Qy      61 CATCTCATGACCATTTGGATTACCCCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 120
Db      579 CATCTCATGACCATTTGCTGCTCATGCTCATCCAGTGTGATCTTCAGGCTATGGATA 638

Qy      121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180
Db      639 GATCAGG-----AGGCATCACTGAGGCCAGGAGCTCTGGCCCATGACCTGTAT 685

Qy      181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGACAGCTTGGGGCTGTGGCCCTGCCCCCT 240
Db      686 CCCAGCTACTCCAACTTCCCATTTCTCGCCCTG-----CCCC 721

Qy      241 TGGTCTCGCCCTAGATACAGACAGTGTATATCCACACACACCTGTCTACAGTGTCTTCAAT 300
Db      722 CGGAGCGAGTCTGTATACGCCCTTTATCTCACAGCTTTTCTACAAATGGCATTCAAT 781

Qy      301 AAAGTGCACGTGTTGTG 318
Db      782 AAAGTGCACGTGTTTCTG 799

RESULT 11
US-09-461-912A-2
; Sequence 2, Application US/09461912A
; Patent No. 6709855
; GENERAL INFORMATION:
; APPLICANT: Stanton, Lawrence A.
; APPLICANT: White, R. Tyler
; APPLICANT: Damm, Deborah L.
```

```
; APPLICANT: Lewicki, John A.
; TITLE OF INVENTION: Methods for detection and use of
; FILE REFERENCE: SCIOS.011A
; CURRENT APPLICATION NUMBER: US/09/461.912A
; CURRENT FILING DATE: 1999-12-15
; PRIOR FILING DATE: 1999-12-15
; PRIOR FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-912A-2

Query Match      23.5%; Score 76; DB 3; Length 362;
Best Local Similarity 100.0%; Pred. No. 4.1e-15;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCCCGAGCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db 285 CGGGCCCGAGCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 344

Qy 61 CATCTCATGACCATT 76
Db 345 CATCTCATGACCATT 360

RESULT 12
US-09-461-912A-1
; Sequence 1, Application US/09461912A
; Patent No. 6709855
; GENERAL INFORMATION:
; APPLICANT: Stanton, Lawrence A.
; APPLICANT: White, R. Tyler
; APPLICANT: Damm, Deborah L.
; APPLICANT: Lewicki, John A.
; TITLE OF INVENTION: Methods for detection and use of
; FILE REFERENCE: SCIOS.011A
; CURRENT APPLICATION NUMBER: US/09/461.912A
; CURRENT FILING DATE: 1999-12-15
; PRIOR FILING DATE: 1999-12-15
; PRIOR FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Rattus rattus
US-09-461-912A-1

Query Match      16.3%; Score 52.6; DB 3; Length 362;
Best Local Similarity 81.3%; Pred. No. 2.5e-07;
Matches 61; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 2 GGGGCCCGAGCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 61
Db 286 GGGGCCCGAGCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 345

Qy 62 ATCTCATGACCATT 76
Db 346 ATCTCATGATTCATT 360

RESULT 13
US-10-131-827-2535
; Sequence 2535, Application US/10131827
; Patent No. 6905827
; GENERAL INFORMATION:
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
```

```
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
; FILE REFERENCE: CHRONIC INFLAMMATORY DISEASES
; CURRENT APPLICATION NUMBER: US/10/131,827
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US 10/006,290
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/296,764
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 9090
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2535
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-827-2535

Query Match      15.5%; Score 50; DB 3; Length 50;
Best Local Similarity 100.0%; Pred. No. 6.5e-07;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 250 CCTAGATACAGCAGTTTATACCCACACACCTCTCTACAGTGCATTCAA 299
Db 1 CCTAGATACAGCAGTTTATACCCACACACCTCTCTACAGTGCATTCAA 50

RESULT 14
US-09-007-005-17/c
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

Query Match      10.7%; Score 34.6; DB 3; Length 289;
Best Local Similarity 5.8%; Pred. No. 0.21;
Matches 10; Conservative 78; Mismatches 83; Indels 0; Gaps 0;

Qy 148 ACGGGTTACTAGTACCGCCCATAGCTGCACACCTTTGCACCTCCACCTGTCATGCTGG 207
Db 250 AYAYGAYGYTYTAYCYGYCYAYGYCYTYGYSYNYSYNYSYNYSYNYSYNYSYNY 191

Qy 208 CCTGCACGCTGGGGTGTGGCCCTGGCCCTGGCTGGCTGGCTGGCTGGCTGGCTGG 267
Db 190 NYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNYSYNY 131

Qy 268 ATACCCACACACCTGTCTACAGTGTCTATTCATTAATAAGTCACGCTGCTGTG 318
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 38.4091 Seconds
(without alignments)
4072.612 Million cell updates/sec

Title: US-09-980-046B-5
Perfect score: 88
Sequence: 1 cggagtcgagccctgtatca.....cgtgtttcttgggtacaaaaa 88

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

- Database : Issued Patents NA:*
- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
 - 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
 - 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
 - 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
 - 5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
 - 6: /cgn2_6/ptodata/1/ina/PTUS COMB.seq:*
 - 7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*
 - 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
 - 9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	90.9	642	3	US-09-370-838-119
2	80	90.9	642	3	US-09-854-133-119
3	75.8	86.1	808	3	US-09-023-655-1279
4	75.8	86.1	808	3	US-09-461-912A-25
5	44.8	50.9	50	3	US-10-131-827-2927
6	42.8	48.6	764	3	US-08-991-789A-283
7	42.8	48.6	764	3	US-09-062-451-283
8	42.8	48.6	764	3	US-09-289-198-283
9	42.8	48.6	764	3	US-09-429-755-283
10	42.8	48.6	764	3	US-09-699-295-283
11	42.8	48.6	764	3	US-09-534-825A-283
12	27.4	31.1	25202	3	US-09-949-016-13151
13	26.8	30.5	513	3	US-09-702-705-314
14	26.8	30.5	513	3	US-09-736-457-314
15	26.8	30.5	513	3	US-09-614-124B-314
16	26.8	30.5	513	3	US-09-671-325-314
17	26.8	30.5	513	3	US-09-589-184-314
18	26.8	30.5	513	3	US-09-658-824-314
19	26.8	30.5	513	3	US-10-017-754-314
20	26.8	30.5	513	3	US-09-651-563-314
21	26.8	30.5	513	3	US-09-519-642-314
22	26.8	30.5	601	3	US-09-949-016-148364
23	26.8	30.5	601	3	US-09-949-016-167312
24	26.8	30.5	40237	3	US-09-949-016-16461

25	26.8	30.5	108169	3	US-09-949-016-12898	Sequence 12898, A
26	26.8	30.5	108169	3	US-09-949-016-15907	Sequence 15907, A
27	26.6	30.2	50	3	US-10-131-827-2535	Sequence 2535, Ap
28	26.4	30.0	601	3	US-09-949-016-108834	Sequence 108834, A
29	26.4	30.0	601	3	US-09-949-016-108835	Sequence 108835, A
c 30	26.4	30.0	60137	3	US-09-949-016-14735	Sequence 14735, A
c 31	26.4	30.0	60137	3	US-09-949-016-14912	Sequence 14912, A
32	25.4	28.9	3362	3	US-09-513-505-5	Sequence 5, Appli
33	25.4	28.9	3382	3	US-09-513-505-7	Sequence 7, Appli
34	25.4	28.9	3420	3	US-09-513-505-1	Sequence 1, Appli
35	25.4	28.9	3478	3	US-09-513-505-3	Sequence 3, Appli
36	25.4	28.9	43360	3	US-09-453-702B-206	Sequence 206, App
37	25.4	28.9	43360	3	US-10-114-170-206	Sequence 206, App
38	25.4	28.9	45325	3	US-09-453-702B-261	Sequence 261, App
39	25.4	28.9	45325	3	US-10-114-170-261	Sequence 261, App
c 40	25.2	28.6	1757	3	US-09-484-970B-37	Sequence 37, Appl
c 41	25.2	28.6	101951	3	US-09-949-016-15648	Sequence 15648, A
c 42	25	28.4	601	3	US-09-949-016-171564	Sequence 171564, A
c 43	25	28.4	601	3	US-09-949-016-171565	Sequence 171565, A
c 44	25	28.4	691	3	US-09-949-016-2545	Sequence 2545, Ap
c 45	25	28.4	695	3	US-09-513-999C-3892	Sequence 3892, Ap

ALIGNMENTS

RESULT 1
US-09-370-838-119
; Sequence 119, Application US/09370838
; Patent No. 644425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370.838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285.323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 119
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-370-838-119

Query Match 90.9%; Score 80; DB 3; Length 642;
Best Local Similarity 94.3%; Pred. No. 3.6e-20;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CGGAGTCGAGTCCTGATCAGCCCTTTTCTACAGCCTTTTCTACAAATGGCATTCAAT 60
Db 535 CGGAGCGAGTCCTGATCAGCCCTTTTCTACAGCCTTTTCTACAAATGGCATTCAAT 594
Qy 61 AAATGTCAGCGTGTCTTCTGGTACAAAAA 88
Db 595 AAATGTCAGCGTGTCTTCTGGTACAAAAA 622

RESULT 2
US-09-854-133-119
; Sequence 119, Application US/09854133
; Patent No. 6759508
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;; FILE REFERENCE: 210121.475C10
;; CURRENT APPLICATION NUMBER: US/09/854.133
;; CURRENT FILING DATE: 2001-05-11
;; NUMBER OF SEQ ID NOS: 735
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 119
;; LENGTH: 642
;; TYPE: DNA
;; ORGANISM: Homo sapien
US-09-854-133-119

Query Match 90.9%; Score 80; DB 3; Length 642;
Best Local Similarity 94.3%; Pred. No. 3.6e-20;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCCTCAGCGCTTTTCTACAATGGCATTCAAT 60
|||||
DB 535 CGGAGCGAGTCCTGTATCAGCCCTTTATCTCCTCAGCGCTTTTCTACAATGGCATTCAAT 594
|||||

QY 61 AAAGTCACGCTGTTCTTGTGTACAAAA 88
|||||
DB 595 AAAGTCACGCTGTTCTTGTGTAAAAAA 622
|||||

RESULT 3
US-09-023-655-1279
; Sequence 1279, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/023.655
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1279:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 808 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g311374

US-09-023-655-1279

Query Match 86.1%; Score 75.8; DB 3; Length 808;
Best Local Similarity 97.5%; Pred. No. 1.5e-18;
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCCTCAGCGCTTTTCTACAATGGCATTCAAT 60
|||||
DB 722 CGGAGCGAGTCCTGTATCAGCCCTTTATCTCCTCAGCGCTTTTCTACAATGGCATTCAAT 781
|||||

QY 61 AAAGTCACGCTGTTCTTGT 79
|||||
DB 782 AAAGTCACGCTGTTCTTGTGG 800
|||||

RESULT 4
US-09-461-912A-25
; Sequence 125, Application US/09461912A
; Patent No. 6709855
; GENERAL INFORMATION:
; APPLICANT: Stanton, Lawrence A.
; APPLICANT: White, R. Tyler
; APPLICANT: Demm, Deborah L.
; APPLICANT: Lewicki, John A.
; TITLE OF INVENTION: Methods for detection and use of
; FILE REFERENCE: SCIOS.011A
; CURRENT APPLICATION NUMBER: US/09/461,912A
; PRIOR FILING DATE: 1999-12-15
; PRIOR FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (238)...(639)
US-09-461-912A-25

Query Match 86.1%; Score 75.8; DB 3; Length 808;
Best Local Similarity 97.5%; Pred. No. 1.5e-18;
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCCTCAGCGCTTTTCTACAATGGCATTCAAT 60
|||||
DB 722 CGGAGCGAGTCCTGTATCAGCCCTTTATCTCCTCAGCGCTTTTCTACAATGGCATTCAAT 781
|||||

QY 61 AAAGTCACGCTGTTCTTGT 79
|||||
DB 782 AAAGTCACGCTGTTCTTGTGG 800
|||||

RESULT 5
US-10-131-827-2927
; Sequence 2927, Application US/10131827
; Patent No. 6905827
; GENERAL INFORMATION:
; APPLICANT: Wohlgemuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES
; FILE REFERENCE: 506612000120
; CURRENT APPLICATION NUMBER: US/10/131,827
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US 10/006,290
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/296,764
; PRIOR FILING DATE: 2001-06-08

; NUMBER OF SEQ ID NOS: 9090
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2927
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-827-2927

Query Match 50.9%; Score 44.8; DB 3; Length 50;
Best Local Similarity 95.8%; Pred. No. 2.4e-07;
Matches 46; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 AGTCAGTCGTGATCAGCCCTTTATCTCTACACGCTTTTCTACAATG 51
|||
Db 3 AGCCGAGTCCTGCATCAGCCCTTTATCTCTACACGCTTTTCTACAATG 50

RESULT 6
US-08-991-789A-283/c
; Sequence 283, Application US/08991789A
; Patent No. 6225054
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; Smith, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 292
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,789A
; FILING DATE: 11-Dec-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 283:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 764 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 283:
US-08-991-789A-283

Query Match 48.6%; Score 42.8; DB 3; Length 764;
Best Local Similarity 75.7%; Pred. No. 3.5e-06;
Matches 53; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 19 CAGCCCTTTATCTCAGCGCTTTTCTACATGGAATGAGTGCAGTGTCTT 78
|||
Db 72 CAGCAGTTTATATCCACACACCTGCTACAGTGTCTACATGGAATGAGTGCAGTGTCTT 78
|||
Qy 79 GGTACAAAAA 88
|||
Db 12 AAAAAAAAAA 3
|||

RESULT 7
US-09-062-451-283/c
; Sequence 283, Application US/09062451
; Patent No. 6344550
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; Smith, John M.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 297
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/062,451
; FILING DATE: 04-APR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.419C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 283:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 764 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-062-451-283

Query Match 48.6%; Score 42.8; DB 3; Length 764;
Best Local Similarity 75.7%; Pred. No. 3.5e-06;
Matches 53; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 19 CAGCCCTTTATCTCAGCGCTTTTCTACATGGAATGAGTGCAGTGTCTT 78
|||
Db 72 CAGCAGTTTATATCCACACACCTGCTACAGTGTCTACATGGAATGAGTGCAGTGTCTT 78
|||
Qy 79 GGTACAAAAA 88
|||
Db 12 AAAAAAAAAA 3
|||

RESULT 8
US-09-289-198-283/c
; Sequence 283, Application US/09289198
; Patent No. 6586570
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; Smith, John M.
; APPLICANT: Reed, Steven G.
; APPLICANT: Misher, Lynda
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C3
; CURRENT APPLICATION NUMBER: US/09/289,198
; CURRENT FILING DATE: 1999-04-09
; EARLIER APPLICATION NUMBER: US 09/062,451
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: US 08/991,789
; EARLIER FILING DATE: 1997-12-11

US-08-991-789A-283/c
; Sequence 283, Application US/08991789A
; Patent No. 6225054
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; Smith, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 292
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed IP Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,789A
; FILING DATE: 11-Dec-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 283:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 764 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 283:
US-08-991-789A-283

Query Match 48.6%; Score 42.8; DB 3; Length 764;
Best Local Similarity 75.7%; Pred. No. 3.5e-06;
Matches 53; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 19 CAGCCCTTTATCTCAGCGCTTTTCTACATGGAATGAGTGCAGTGTCTT 78
|||
Db 72 CAGCAGTTTATATCCACACACCTGCTACAGTGTCTACATGGAATGAGTGCAGTGTCTT 78
|||
Qy 79 GGTACAAAAA 88
|||
Db 12 AAAAAAAAAA 3
|||

RESULT 7
US-09-062-451-283/c
; Sequence 283, Application US/09062451
; Patent No. 6344550
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; Smith, John M.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 297
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/062,451
; FILING DATE: 04-APR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.419C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 283:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 764 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-062-451-283

Query Match 48.6%; Score 42.8; DB 3; Length 764;
Best Local Similarity 75.7%; Pred. No. 3.5e-06;
Matches 53; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 19 CAGCCCTTTATCTCAGCGCTTTTCTACATGGAATGAGTGCAGTGTCTT 78
|||
Db 72 CAGCAGTTTATATCCACACACCTGCTACAGTGTCTACATGGAATGAGTGCAGTGTCTT 78
|||
Qy 79 GGTACAAAAA 88
|||
Db 12 AAAAAAAAAA 3
|||

RESULT 8
US-09-289-198-283/c
; Sequence 283, Application US/09289198
; Patent No. 6586570
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; Smith, John M.
; APPLICANT: Reed, Steven G.
; APPLICANT: Misher, Lynda
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C3
; CURRENT APPLICATION NUMBER: US/09/289,198
; CURRENT FILING DATE: 1999-04-09
; EARLIER APPLICATION NUMBER: US 09/062,451
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: US 08/991,789
; EARLIER FILING DATE: 1997-12-11


```
Query Match      48.6%; Score 42.8; DB 3; Length 764;
Best Local Similarity 75.7%; Pred. No. 3.5e-06;
Matches 53; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 19 CAGCCCTTATCCCTCACACGGCTTTCTACAAATGGCATTCAATAAAGTGCAGGTGTTCTT 78
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 72 CAGCAGTTTATATACCCACACACCTGTCTACAGTGTCTATCAATAAAGTGCAGGTGCTTGGC 13
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 79 GGTACAAAAA 88
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12 AAAAAAANA 3
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
US-09-949-016-13151
; Sequence 13151, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13151
; LENGTH: 25202
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13151

Query Match      31.1%; Score 27.4; DB 3; Length 25202;
Best Local Similarity 59.7%; Pred. No. 6.9;
Matches 46; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 6 TCGAGTCCTGTATCAGCCCTTTATCCTCACACGGCTTTCTACAAATGGCATTCAATAAAGT 65
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 TAGAATCCTGAGTCAGCTTTTCCCTCCTACTCTGATTCCTACCCCTTAATCAAAAT 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 66 GCACGTGTTCTTGTA 82
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 TCCATTGTTGGTGCA 198
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-09-702-705-314/c
; Sequence 314, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 314
```

```
; LENGTH: 513
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-314

Query Match      30.5%; Score 26.8; DB 3; Length 513;
Best Local Similarity 64.5%; Pred. No. 3;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 26 TTATCCTCACACGGCTTTTCTACAATGGCATTCAATAAAGTGCAGGTGTTCTTGGTACAA 85
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 198 TTTACCTTTCACGTCTTCTATATGTTATTAATTAAGTTACGCCATTTATGGGGAGAA 139
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 86 AA 87
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 138 AA 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
US-09-736-457-314/c
; Sequence 314, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 314
; LENGTH: 513
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-314

Query Match      30.5%; Score 26.8; DB 3; Length 513;
Best Local Similarity 64.5%; Pred. No. 3;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 26 TTATCCTCACACGGCTTTTCTACAATGGCATTCAATAAAGTGCAGGTGTTCTTGGTACAA 85
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 198 TTTACCTTTCACGTCTTCTATATGTTATTAATTAAGTTACGCCATTTATGGGGAGAA 139
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 86 AA 87
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 138 AA 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
US-09-614-124B-314/c
; Sequence 314, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
```

; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 314
; LENGTH: 513
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-614-124B-314

Query Match 30.5%; Score 26.8; DB 3; Length 513;
Best Local Similarity 64.5%; Pred. No. 3;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Qy 26 TTATCCTCACACGCTTTTCTACAAATGGCATTCAATAAAGTGCACGCTTTCTTGGTACAA 85
Db 198 TTTACCTTTCACCTGATTTCTATATGGTATAAATTAAGTTCAGGCATTTTATGGGGAGAA 139
Qy 86 AA 87
Db 138 AA 137

Search completed: December 6, 2005, 11:47:20
Job time : 40.4091 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 787.105 Seconds
(without alignments)
6427.437 Million cell updates/sec

Title: US-09-980-046B-4
Perfect score: 89
Sequence: 1 cgggaatcccccttcagttc.....aaaaaaaaaaaaaagaaaaa 89

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_by.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79.4	89.2	1464	BD270355	BD270355 50 human
2	79.4	89.2	1757	AK223104	AK223104 Homo sapi
3	70	78.7	428	CQ923987	CQ923987 Sequence
4	70	78.7	522	CS113124	CS113124 Sequence
5	70	78.7	592	AX780236	AX780236 Sequence
6	70	78.7	1761	BC017716	BC017716 Homo sapi
7	70	78.7	1766	AK223068	AK223068 Homo sapi
8	70	78.7	2000	CQ776411	CQ776411 Sequence
9	70	78.7	2000	CQ861464	CQ861464 Sequence
10	70	78.7	2000	CQ981630	CQ981630 Sequence
11	70	78.7	2000	AR569106	AR569106 Sequence
12	70	78.7	2000	AX055560	AX055560 Sequence
13	70	78.7	2000	AX333755	AX333755 Sequence
14	70	78.7	2000	HSU77643	U77643 Homo sapien
15	70	78.7	2180	AR071760	AR071760 Sequence
16	70	78.7	2180	AR097657	AR097657 Sequence
17	70	78.7	2180	AR195192	AR195192 Sequence
18	70	78.7	71352	AC137090	AC137090 Homo sapi

19	70	78.7	154252	8	AC132938	AC132938 Homo sapi
20	70	78.7	166458	14	AC084737	AC084737 Homo sapi
21	70	78.7	171569	8	AC132872	AC132872 Homo sapi
22	54	60.7	401	6	AX778685	AX778685 Sequence
23	43.6	49.0	177	6	BD112129	BD112129 EST and e
24	43.6	49.0	177	6	AR416576	AR416576 Sequence
25	43.6	49.0	177	6	AX977270	AX977270 Sequence
26	43.6	49.0	161296	2	AC091483	AC091483 Trypanoso
27	43.2	48.5	2720	9	BC058402	BC058402 Mus muscu
28	42.6	47.9	208462	9	AC129179	AC129179 Mus muscu
29	42.6	47.9	232320	9	AC127314	AC127314 Mus muscu
30	42.6	47.9	349980	6	AX344565	AX344565 Sequence
31	42.2	47.4	1329	8	BC053641	BC053641 Homo sapi
32	42.2	47.4	1386	5	BC090110	BC090110 xenopus t
33	42.2	47.4	1526	8	BC041705	BC041705 Homo sapi
34	41.6	46.7	1647	8	AF218005	AF218005 Homo sapi
35	41.6	46.7	1961	8	AB060916	AB060916 Macaca fa
36	41.6	46.7	3273	9	BC023693	BC023693 Mus muscu
37	41.6	46.7	349980	6	AX344570	AX344570 Sequence
38	41.4	46.5	1817	9	BC051174	BC051174 Mus muscu
39	41.4	46.5	2057	8	BC017236	BC017236 Homo sapi
40	41.4	46.5	110000	14	PFMAL13_24	Continuation (25 o
41	41.4	46.5	138622	9	AC124401	AC124401 Mus muscu
42	41.4	46.5	179155	9	AL772401	AL772401 Mouse DNA
43	41.4	46.5	208458	9	AC118009	AC118009 Mus muscu
44	41.4	46.5	219132	9	AC151478	AC151478 Mus muscu
45	41.2	46.3	216215	8	HSG256022	AL080239 Human DNA

ALIGNMENTS

RESULT 1	LOCUS	BD270355	50 human secreted proteins.	1464 bp	DNA	linear	PAT 17-JUL-2003
BD270355	DEFINITION	BD270355	50 human secreted proteins.				
	ACCESSION	BD270355					
	VERSION	BD270355.1	GI:33080123				
	KEYWORDS	JP 2002539775-A/44.					
	SOURCE	Homo sapiens (human)					
	ORGANISM	Homo sapiens					
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.					
	AUTHORS	1 (bases 1 to 1464)					
	TITLE	Rosen,C.A., Ruben,S.M. and Komatsoulis,G.					
	JOURNAL	50 human secreted proteins					
	COMMENT	Patent: JP 2002539775-A 44 26-NOV-2002; HUMAN GENOME SCIENCES INC					
		OS Homo sapiens (human)					
		PN JP 2002539775-A/44					
		PD 26-NOV-2002					
		PF 09-MAR-2000 JP 2000606612					
		PR 19-WAR-1999 US 60/125360,11-JUN-1999 US 60/138626 PR					
		03-DEC-1999 US 60/168662					
		PI CRAIG A ROSEN,STEVEN M RUBEN,GEORGE KOMATSOUKIS PC					
		C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P1/00,A61P1/14, PC					
		A61P1/16,					
		PC					
		A61P3/00,A61P3/06,A61P3/10,A61P7/00,A61P7/02,A61P7/06,A61P9/00, PC					
		A61P9/06,					
		PC A61P9/10,A61P9/10,A61P15/00,A61P17/02,A61P17/06,A61P19/00, PC					
		A61P25/00,					
		PC A61P25/14,A61P25/16,A61P25/24,A61P25/28,A61P27/00,A61P27/06,					
		PC A61P29/00,					
		PC A61P31/00,A61P31/04,A61P31/14,A61P31/16,A61P31/18,A61P31/20,					
		PC A61P31/22,					
		PC A61P33/00,A61P33/02,A61P33/06,A61P35/00,A61P35/02,A61P37/00,					
		PC A61P37/06,					
		PC A61P37/08,A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC					
		C12N1/21,					
		PC C12N5/10,C12P21/02,C12Q1/02,C12Q1/68,G01N33/68// (C12P21/02, PC					
		C12R1:19),					

```

PC      C12N15/00.A61K37/02.C12N5/00
CC      50 human secreted proteins
FH      Key      Location/Qualifiers
FT      source      1..1464
FT      Location/Qualifiers
FEATURES             source      1..1464
                             /organism="Homo sapiens (human)"
                             /mol_type="genomic DNA"
                             /db_xref="taxon:9606"
ORIGIN
Query Match      89.2%; Score 79.4; DB 6; Length 1464;
Best Local Similarity 93.3%; Pred. No. 1.2e-07;
Matches 83; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy      1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db      1071 CGGGAATCCCCCTTCAGTCTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1130
Qy      61 AACAAACCAAAAAAAAAAAAAAAAAAAAAA 89
Db      1131 AACAAACCGACTCACAAAAAAAAAAAAAAAA 1159
RESULT 2
LOCUS      AK223104      1757 bp      mRNA      linear      PRI 26-APR-2005
DEFINITION      Homo sapiens mRNA for secreted and transmembrane 1 precursor
ACCESSION      AK223104
VERSION      AK223104.1 GI:62897768
KEYWORDS      FLI_CDNA; oligo capping.
SOURCE      Homo sapiens
ORGANISM      Homo sapiens
              Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominiidae; Homo.
REFERENCE      1
AUTHORS      Maruyama,K. and Sugano,S.
TITLE      Oligo-capping: a simple method to replace the cap structure of
              eukaryotic mRNAs with oligoribonucleotides
JOURNAL      Gene 138 (1-2), 171-174 (1994)
PUBMED      8125298
REFERENCE      2
AUTHORS      Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
              Sugano,S.
TITLE      Construction and characterization of a full length-enriched and a
              5'-end-enriched cDNA library
JOURNAL      Gene 200 (1-2), 149-156 (1997)
PUBMED      9373149
REFERENCE      3 (bases 1 to 1757)
AUTHORS      Suzuki,Y., Sugano,S., Totoki,Y., Toyoda,A., Takeda,T., Sakaki,Y.,
              Tanaka,A. and Yokoyama,S.
TITLE      Direct Submission
JOURNAL      Submitted (22-APR-2005) Akiko Tanaka, RIKEN Yokohama Institute,
              Protein Research Group; 1-7-22 Suehiro, Tsurumi, Yokohama,
              Kanagawa, 230-0045, Japan (E-mail:aktanaka@riken.jp,
              URL:http://protein.gsc.riken.jp/, Tel:81-45-503-9452,
              Fax:81-45-503-9450)
COMMENT      This work was supported in part by the National Project on Protein
              Structural and Functional Analysis, Ministry of Education, Culture,
              Sports, Science and Technology of Japan.
              Sumio Sugano, Yutaka Suzuki
              Laboratory of Functional Genomics Department of Medical Genome
              Sciences Graduate School of Frontier Sciences The University of
              Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639 Japan email:
              sugano@k.u-tokyo.ac.jp
              URL:http://www.k.u-tokyo.ac.jp/index.html.en.
              Location/Qualifiers
              1..1757
              /organism="Homo sapiens"
              /mol_type="mRNA"
FEATURES             source
              Location/Qualifiers

```

```

/db_xref="taxon:9606"
/clone="KAT04885"
/cell_line="KATO III"
/clone_lib="KAT"
/notes="cloning vector: pME18SFL3
this clone is also named as hss001002732"
<123..869
/note="Start codon is not identified."
/codon_start=1
/evidence=not_experimental
/product="secreted and transmembrane 1 precursor variant"
/protein_id="BAD96824.1"
/db_xref="GI:62897769"
/tranlations="MOTCLAPPGHVSQALGTLFLAASLSAONEGWDSPICTEGVVS
VSGENTVSCNISAFSHVNIKLRAHQBSAIFNEVAPGYFDFRQWLQVQGGVAQL
VIKGADSHAGLHYMHLVGHQRNNRQVTLVEVSAGBPSADPTGFWPVPVAVTAVILL
VALVFWYRCRSQORREKFFLEPQMKVAALRAGAAQQGLSRASAEALWTPDSEPTP
RPLALVFKPSPLGALLSPQLFPYAADP"
CDS
Query Match      89.2%; Score 79.4; DB 8; Length 1757;
Best Local Similarity 93.3%; Pred. No. 1.1e-07;
Matches 83; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy      1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db      1650 CGGGAATCCCCCTTCAGTCTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1709
Qy      61 AACAAACCAAAAAAAAAAAAAAAAAAAAAA 89
Db      1710 AACAAACCGACTCACAAAAAAAAAAAAAAAA 1738
RESULT 3
LOCUS      CQ923987      428 bp      DNA      linear      PAT 23-NOV-2004
DEFINITION      Sequence 5187 from Patent WO2004097052.
ACCESSION      CQ923987
VERSION      CQ923987.1 GI:56213928
KEYWORDS      .
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominiidae; Homo.
REFERENCE      1
AUTHORS      Burczynski,M.E., Twine,N.C., Slonim,D.K., Trepicchio,W.L.,
              Strahs,A., Immerman,F. and Dörner,A.J.
TITLE      Methods for prognosis and treatment of solid tumors
JOURNAL      Patent: WO 2004097052-A 5187 11-NOV-2004;
              Wyeth (US); Burczynski, Michael E. (US)
FEATURES             source      1..428
              Location/Qualifiers
              1..428
              /organism="Homo sapiens"
              /mol_type="unassigned DNA"
              /db_xref="taxon:9606"
              misc_feature      234..240
              /notes="n includes a, c, g, or t, or contains no
              nucleotide"
              misc_feature      242..249
              /notes="n includes a, c, g, or t, or contains no
              nucleotide"
              misc_feature      252..256
              /notes="n includes a, c, g, or t, or contains no
              nucleotide"
ORIGIN
Query Match      78.7%; Score 70; DB 6; Length 428;
Best Local Similarity 93.6%; Pred. No. 1.8e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy      1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60

```

Db 100 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 159

Qy 61 AACAAACCAAAAAA 78
 |||||
 Db 160 AACAAACCGACTCACA 177

RESULT 4
 CS113124/c 522 bp DNA linear PAT 24-JUN-2005
 DEFINITION Sequence 142 from Patent WO2005054507.
 ACCESSION CS113124
 VERSION CS113124.1 GI:68224696
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Corfe, B. and Chirakkal, H.
 TITLE Gene screen
 JOURNAL Patent: WO 2005054507-A 142 16-JUN-2005;
 University of Sheffield (GB)
 FEATURES
 source Location/Qualifiers
 1..522
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 78.7%; Score 70; DB 6; Length 522;
 Best Local Similarity 93.6%; Pred. No. 1.7e-05;
 Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 Db 82 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 23
 |||||
 Qy 61 AACAAACCAAAAAA 78
 |||||
 Db 22 AACAAACCGACTCACA 5

RESULT 5
 AX780236 592 bp DNA linear PAT 14-JUL-2003
 LOCUS Sequence 2393 from Patent WO03039443.
 DEFINITION AX780236
 ACCESSION AX780236
 VERSION AX780236.1 GI:32697230
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Haeflrich, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,
 Dugas, M., Eils, R., Brors, B. and Wergenthaler, S.
 TITLE Novel genetic markers for leukemias
 JOURNAL Patent: WO 03039443-A 2393 15-MAY-2003;
 Deutsches Krebsforschungszentrum (DE) ;
 Ludwig-Maximilian-Universitaet Muenchen (DE) ;
 PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)
 FEATURES
 source Location/Qualifiers
 1..592
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 78.7%; Score 70; DB 6; Length 592;
 Best Local Similarity 93.6%; Pred. No. 1.7e-05;

Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 Db 434 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 493
 |||||

Qy 61 AACAAACCAAAAAA 78
 |||||
 Db 494 AACAAACCGACTCACA 511

RESULT 6

BC017716

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 1761)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.W., Hong, L.,

Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,

Villalón, D.C., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,

Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,

Schneerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1761)

Strausberg, R.

Direct Submission

Submitted (03-DEC-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-sbgc.stanford.edu>

Contact: (Dickson, Mark) mcdepaxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IPAK Plate: 27 Row: h Column: 10

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 4506868.

Location/Qualifiers

```
source
1. .1761
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:21265 IMAGE:4413387"
/tissue_type="Liver, adenocarcinoma"
/clone_lib="NIH_MGC_90"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1. .1761
/genes="SECTM1"
/notes="synonym: K12"
/db_xref="GeneID:6398"
/db_xref="MIM:602602"
93. .839
/genes="SECTM1"
/codon_start=1
/product="secreted and transmembrane 1 precursor"
/protein_id="AAH17716.1"
/db_xref="GI:17389346"
/db_xref="GeneID:6398"
/db_xref="MIM:602602"
/translaton="MOTCLAPFGHVSOALGTLFLAASLSAQNEGWDSPICTEGVVS
VSGENTVMSCNISNAFNVHNIKLRAHQESAFNEVAPGYFRDGHQLOVQGVQVL
VIKGARDSHAGLVMHVLVGHORNNRQVTLVSGAEPOSADPTGFWPVPVAVTVFILL
VALVMPAWYRCRCQQRRKKFFLLBPQMKFAALRAGAQQGLSRASALRTPDSEPTP
RPLALVFKPSPLGALLELLSPQLFPYAADP"

gene
1. .1761
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KAT02983"
/cell_line="KATO III"
/clone_lib="KAT"
/notes="cloning vector: pME18SFL3
this clone is also named as hss001002587"
<92. .838
/notes="Start codon is not identified."
/codon_start=1
/evidence="not_experimental
/product="secreted and transmembrane 1 precursor variant"
/protein_id="BAD96788.1"
/db_xref="GI:62897697"
/translaton="MOTCLAPFGHVSOALGTLFLAASLSAQNEGWDSPICTEGVVS
VSGENTVMSCNISNAFNVHNIKLRAHQESAFNEVAPGYFRDGHQLOVQGVQVL
VIKGARDSHAGLVMHVLVGHORNNRQVTLVSGAEPOSADPTGFWPVPVAVTVFILL
VALVMPAWYRCRCQQRRKKFFLLBPQMKFAALRAGAQQGLSRASALRTPDSEPTP
RPLALVFKPSPLGALLELLSPQLFPYAADP"

CDS
Query Match 78.7%; Score 70; DB 8; Length 1761;
Best Local Similarity 93.6%; Pred. No. 1.2e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
Db 1620 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 1679

Qy 61 AACAAACCAAAAAA 78
Db 1680 AACAAACCGACTCAGAA 1697

RESULT 7
AK223068 LOCUS Homo sapiens mRNA for secreted and transmembrane 1 precursor
DEFINITION variant, clone: KAT02983.
ACCESSION AK223068
VERSION AK223068.1 GI:62897696
KEYWORDS FLI_CDNA; oligo capping.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 Maruyama,K. and Sugano,S.
AUTHORS Oligo-capping: a simple method to replace the cap structure of
TITLE eukaryotic mRNAs with oligoribonucleotides
JOURNAL Gene 138 (1-2), 171-174 (1994)
PUBMED 8125298
REFERENCE
2 Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
AUTHORS Sugano,S.
TITLE Construction and characterization of a full length-enriched and a
5'-end-enriched cDNA library
JOURNAL Gene 200 (1-2), 149-156 (1997)
PUBMED 9373149
REFERENCE
3 (bases 1 to 1766)
AUTHORS Suzuki,Y., Sugano,S., Totoki,Y., Toyoda,A., Takeda,T., Sakaki,Y.,
Tanaka,A. and Yokoyama,S.
TITLE Direct Submission
JOURNAL Submitted (22-APR-2005) Akiko Tanaka, RIKEN Yokohama Institute,
```

```
Protein Research Group; 1-7-22 Suehiro, Teurumi, Yokohama,
Kanagawa, 230-0045, Japan (E-mail:aktanaka@riken.jp,
URL:http://protein.gsc.riken.jp/, tel:81-45-503-9452,
Fax:81-45-503-9450)
This work was supported in part by the National Project on Protein
Structural and Functional Analysis, Ministry of Education, Culture,
Sports, Science and Technology of Japan.
Sumio Sugano, Yutaka Suzuki
Laboratory of Functional Genomics Department of Medical Genome
Sciences Graduate School of Frontier Sciences The University of
Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639 Japan email:
sugano@k.u-tokyo.ac.jp
URL: http://www.k.u-tokyo.ac.jp/index.html.en.

FEATURES
source
1. .1766
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KAT02983"
/cell_line="KATO III"
/clone_lib="KAT"
/notes="cloning vector: pME18SFL3
this clone is also named as hss001002587"
<92. .838
/notes="Start codon is not identified."
/codon_start=1
/evidence="not_experimental
/product="secreted and transmembrane 1 precursor variant"
/protein_id="BAD96788.1"
/db_xref="GI:62897697"
/translaton="MOTCLAPFGHVSOALGTLFLAASLSAQNEGWDSPICTEGVVS
VSGENTVMSCNISNAFNVHNIKLRAHQESAFNEVAPGYFRDGHQLOVQGVQVL
VIKGARDSHAGLVMHVLVGHORNNRQVTLVSGAEPOSADPTGFWPVPVAVTVFILL
VALVMPAWYRCRCQQRRKKFFLLBPQMKFAALRAGAQQGLSRASALRTPDSEPTP
RPLALVFKPSPLGALLELLSPQLFPYAADP"

ORIGIN
Query Match 78.7%; Score 70; DB 8; Length 1766;
Best Local Similarity 93.6%; Pred. No. 1.2e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
Db 1619 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 1678

Qy 61 AACAAACCAAAAAA 78
Db 1679 AACAAACCGACTCAGAA 1696

RESULT 8
CQ776411 LOCUS Homo sapiens
DEFINITION Sequence 97 from Patent EP1394274.
ACCESSION CQ776411
VERSION CQ776411.1 GI:45379801
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 Otani,N., Sugita,Y., Yamaya,M., Kubo,H., Nagai,H. and Izuwara,K.
AUTHORS Methods of testing for bronchial asthma or chronic obstructive
TITLE pulmonary disease
JOURNAL Patent: EP 1394274-A 97 03-MAR-2004;
PUBMED Genex Research, Inc. (JP)
FEATURES Location/Qualifiers
source 1. .2000
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

ORIGIN	source	1. .2000	/organism="Homo sapiens"	/mol_type="unassigned DNA"	/db_xref="taxon:9606"
Query Match	78.7%;	Score 70;	DB 6;	Length 2000;	
Best Local Similarity	93.6%;	Pred. No. 1.1e-05;			
Matches	73;	Conservative	0;	Mismatches	5;
Indels	0;	Gaps	0;		
QY	1	CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCATGCTCGAATATCTGAAATGAAGAA	60		
Db	1645	CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCATGCTCGAATATCTGAAATGAAGAA	1704		
QY	61	AACAAACCAAAAAA 78			
Db	1705	AACAAACCGACTCACAA 1722			
RESULT 9					
LOCUS	Q0861464	2000 bp	DNA	linear	PAT 10-SEP-2004
DEFINITION	Sequence 97 from Patent WO2004072265.				
ACCESSION	Q0861464				
VERSION	Q0861464.1	GI:51982453			
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1	Burczynski, M., Twine, N., Dörner, A. J. and Trepicchio, W. L.			
AUTHORS	METHODS FOR MONITORING DRUG ACTIVITIES IN VIVO / I				
TITLE	Patent: WO 2004072265-A 97 26-AUG-2004;				
JOURNAL	Wyeth (US); Burczynski, Michael E. (US); Twine, Natalie C. (US); Dörner, Andrew J. (US); Trepicchio, William L. (US)				
FEATURES	Location/Qualifiers				
source	1. .2000				
/organism="Homo sapiens"	/mol_type="unassigned DNA"				
/db_xref="taxon:9606"					
ORIGIN					
Query Match	78.7%;	Score 70;	DB 6;	Length 2000;	
Best Local Similarity	93.6%;	Pred. No. 1.1e-05;			
Matches	73;	Conservative	0;	Mismatches	5;
Indels	0;	Gaps	0;		
QY	1	CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCATGCTCGAATATCTGAAATGAAGAA	60		
Db	1645	CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCATGCTCGAATATCTGAAATGAAGAA	1704		
QY	61	AACAAACCAAAAAA 78			
Db	1705	AACAAACCGACTCACAA 1722			
RESULT 10					
LOCUS	Q0981630	2000 bp	DNA	linear	PAT 25-JAN-2005
DEFINITION	Sequence 485 from Patent EPI498424.				
ACCESSION	Q0981630				
VERSION	Q0981630.1	GI:58190920			
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1	Rosenthal, A., Hermann, K., Heiden, E., Pilarsky, C., Bruemendorf, T., Straub, E., Roepcke, S., Mennerich, D., Kinnemann, H. and Li, X.			
AUTHORS	Human nucleic acid sequences from lung tumors				
TITLE	Patent: EP 1498424-A 485 19-JAN-2005;				
JOURNAL	Hinzmann, Bernd (DE); Hermann, Klaus (DE); Heiden, Esméralda (DE); Rosenthal, Andre (DE)				
FEATURES	Location/Qualifiers				
source	1. .2000				
/organism="Homo sapiens"	/mol_type="unassigned DNA"				
/db_xref="taxon:9606"					
ORIGIN					
Query Match	78.7%;	Score 70;	DB 6;	Length 2000;	
Best Local Similarity	93.6%;	Pred. No. 1.1e-05;			
Matches	73;	Conservative	0;	Mismatches	5;
Indels	0;	Gaps	0;		
QY	1	CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCATGCTCGAATATCTGAAATGAAGAA	60		
Db	1645	CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCATGCTCGAATATCTGAAATGAAGAA	1704		
QY	61	AACAAACCAAAAAA 78			
Db	1705	AACAAACCGACTCACAA 1722			
RESULT 11					
LOCUS	AR569106	2000 bp	DNA	linear	PAT 14-DEC-2004
DEFINITION	Sequence 3 from patent US 6762030.				
ACCESSION	AR569106				
VERSION	AR569106.1	GI:56569141			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 2000)				
AUTHORS	Lyman, S. D. and Fanslow, W. C. III.				
TITLE	Ligand for CD7 and methods for use thereof				
JOURNAL	Patent: US 6762030-A 3 13-JUL-2004; Immunex Corporation; Seattle, WA; WOX;				
FEATURES	Location/Qualifiers				
source	1. .2000				
/organism="unknown"	/mol_type="genomic DNA"				
ORIGIN					
Query Match	78.7%;	Score 70;	DB 6;	Length 2000;	
Best Local Similarity	93.6%;	Pred. No. 1.1e-05;			
Matches	73;	Conservative	0;	Mismatches	5;
Indels	0;	Gaps	0;		
QY	1	CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCATGCTCGAATATCTGAAATGAAGAA	60		
Db	1645	CGGGAAT			

```
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
119..865
/notes="unnamed protein product"
/codon_start=1
/db_xref="GI:12228779"
/translation="MQTCLAPFGHVSQALGTLFLAASLSAONEGWDSPICTEGVVS
VSGENTVMSCNISAFSHVNIKLRAHQESALFNEVAPGYFRDGMQLQVQGGVAQL
VIKGRDASHAGLYMHVGHQRNNRQVTLVSGAEQPSADPTGFWPVPVAVTAVFILL
VALVMPAWYRCRCQQRREKFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTP
RPLALVPKPSPLGALLELSQPPLFFPYADP"

ORIGIN
Query Match 78.7%; Score 70; DB 6; Length 2000;
Best Local Similarity 93.6%; Pred. No. 1.1e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 60
Dy 1645 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Dy 1705 AACAAACCGACTCACA 1722

RESULT 13
AX333755 2000 bp DNA linear PAT 09-JAN-2002
LOCUS AX333755 Sequence 4264 from Patent WO0194629.
DEFINITION AX333755
ACCESSION AX333755
VERSION AX333755.1 GI:18124474
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrihan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 4264 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source
1..2000
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 78.7%; Score 70; DB 6; Length 2000;
Best Local Similarity 93.6%; Pred. No. 1.1e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 60
Dy 1645 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Dy 1705 AACAAACCGACTCACA 1722

RESULT 14
HSU77643 2000 bp mRNA linear PRI 10-MAR-1998
LOCUS HSU77643
DEFINITION Homo sapiens K12 protein precursor mRNA, complete cds.
AUTHORS U77643
ACCESSION U77643
VERSION U77643.1 GI:2062390
```

```
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2000)
AUTHORS Slentz-Kesler, K.A., Hale, L.P. and Kaufman, R.E.
TITLE Identification and characterization of K12 (SECTM1), a novel human
gene that encodes a Golgi-associated protein with transmembrane and
secreted isoforms
JOURNAL Genomics 47 (3), 327-340 (1998)
PUBMED 9480746
REFERENCE 2 (bases 1 to 2000)
AUTHORS Slentz-Kesler, K.A. and Kaufman, R.E.
TITLE Direct Submission
JOURNAL DUMC, Durham, NC 27710, USA
SUBMITTED (07-NOV-1996) Biochemistry, Duke University, Box 3250
DUMC, Durham, NC 27710, USA
FEATURES
source
1..2000
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q25.2-25.3"
/cell_type="K562 erythroleukemic"
119..865
/notes="type Ia transmembrane protein"
/codon_start=1
/product="K12 protein precursor"
/protein_id="AAC52044.1"
/db_xref="GI:2062391"
/translation="MQTCLAPFGHVSQALGTLFLAASLSAONEGWDSPICTEGVVS
VSGENTVMSCNISAFSHVNIKLRAHQESALFNEVAPGYFRDGMQLQVQGGVAQL
VIKGRDASHAGLYMHVGHQRNNRQVTLVSGAEQPSADPTGFWPVPVAVTAVFILL
VALVMPAWYRCRCQQRREKFFLEPQMKVAALRAGAQQGLSRASAEIWTDPSEPTP
RPLALVPKPSPLGALLELSQPPLFFPYADP"
119..202
sig_peptide
203..862
mat_peptide
551..623
misc_feature
/notes="encodes transmembrane domain"
752..862
misc_feature
/notes="encodes proline-rich region"

ORIGIN
Query Match 78.7%; Score 70; DB 8; Length 2000;
Best Local Similarity 93.6%; Pred. No. 1.1e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 60
Dy 1645 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Dy 1705 AACAAACCGACTCACA 1722

RESULT 15
AR071760 2180 bp DNA linear PAT 18-FEB-2000
LOCUS AR071760 Sequence 2 from patent US 5912142.
DEFINITION AR071760
ACCESSION AR071760
VERSION AR071760.1 GI:7222648
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2180)
AUTHORS Kaufman, R.E. and Slentz-Kesler, K.
TITLE Gene product over expressed in cancer cells
JOURNAL Patent: US 5912142-A 2 15-JUN-1999;
```

FEATURES
source Location/Qualifiers
1..2180
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 78.7%; Score 70; DB 6; Length 2180;
Best Local Similarity 93.6%; Pred. No. 1.1e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCGATGACTCGAATATCTGAATGAAGAA 60
|||
Db 1841 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCGATGACTCGAATATCTGAATGAAGAA 1900
|||

Qy 61 AACAAACCAAAAAAAA 78
|||
Db 1901 AACAAACCGACTCACAA 1918
|||

Search completed: December 7, 2005, 04:21:58
Job time : 789.105 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 698.666 Seconds
(without alignments)
6427.437 Million cell updates/sec

Title: US-09-980-046B-3
Perfect score: 79
Sequence: 1 cgggaatcccccttcagtc.....aaacaaaccgactctaaaaa 79

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb ba.*
- 2: gb in.*
- 3: gb env.*
- 4: gb on.*
- 5: gb ov.*
- 6: gb pat.*
- 7: gb ph.*
- 8: gb pr.*
- 9: gb ro.*
- 10: gb sts.*
- 11: gb sy.*
- 12: gb un.*
- 13: gb vl.*
- 14: gb htg.*
- 15: gb pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75.8	95.9	1464	6	BD270355 50 human
2	75.8	95.9	1757	8	AK223104 Homo sapi
3	74.8	94.7	428	6	CQ923987 Sequence
c	74.8	94.7	522	6	CS113124 Sequence
	74.8	94.7	592	6	AX780236 Sequence
5	74.8	94.7	1761	8	BC017716 Homo sapi
6	74.8	94.7	1766	8	AK223068 Homo sapi
7	74.8	94.7	2000	6	CQ776411 Sequence
8	74.8	94.7	2000	6	CQ861464 Sequence
9	74.8	94.7	2000	6	CQ981630 Sequence
10	74.8	94.7	2000	6	AR569106 Sequence
11	74.8	94.7	2000	6	AX055560 Sequence
12	74.8	94.7	2000	6	AX333755 Sequence
13	74.8	94.7	2000	8	HSU77643
14	74.8	94.7	2180	6	AR071760 Sequence
15	74.8	94.7	2180	6	AR097657 Sequence
16	74.8	94.7	2180	6	AR195192 Sequence
17	74.8	94.7	71352	14	AC137090 Homo sapi
18	74.8	94.7	71352	14	AC137090 Homo sapi

74.8	94.7	154252	8	AC132938	AC132938	Homo sapi
74.8	94.7	166458	14	AC084737	AC084737	Homo sapi
74.8	94.7	171569	8	AC132872	AC132872	Homo sapi
54	68.4	401	6	AX778685	Sequence	
34.2	43.3	174294	14	AC013659	AC013659	Homo sapi
34.2	43.3	186783	8	AC090355	AC090355	Homo sapi
34.2	43.3	187167	14	AP001563	AP001563	Homo sapi
33.8	42.8	161296	2	AC091483	Trypanoso	
33.4	42.3	178524	8	AC055120	Homo sapi	
33.4	42.3	191840	8	AC092656	AC092656	Homo sapi
32.2	40.8	65672	14	AC100892	Mus muscu	
32.2	40.8	152306	9	AC132290	Mus muscu	
32.2	40.8	226007	14	AC153360	Mus muscu	
32.2	40.8	255667	9	AC161426	Mus muscu	
31.8	40.3	205793	14	AC126989	Rattus no	
31.8	40.3	224158	14	AC093984	Rattus no	
31.8	40.3	234778	9	AC115778	Mus muscu	
31.8	40.3	277424	14	AC120729	Rattus no	
31.2	39.5	167585	8	AL137017	Human DNA	
30.8	39.0	208462	9	AC129179	Mus muscu	
30.8	39.0	232320	9	AC127314	Mus muscu	
30.6	38.7	162281	8	AC105395	Homo sapi	
30.6	38.7	189631	14	AC009920	Homo sapi	
30.6	38.7	191935	14	AC114747	Homo sapi	
30.6	38.7	196318	14	AP001950	Homo sapi	
30.6	38.7	203114	14	AC011818	Homo sapi	
30.6	38.7	206364	14	CT009494	Mus muscu	

ALIGNMENTS

RESULT 1	BD270355	50 human secreted proteins.	1464 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD270355	50 human secreted proteins.				
DEFINITION	BD270355	50 human secreted proteins.				
ACCESSION	BD270355	50 human secreted proteins.				
VERSION	BD270355.1	GI:33080123				
KEYWORDS	JP 2002539775-A/44.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 1464)					
AUTHORS	Rosen,C.A., Ruben,S.M. and Komateoulis,G.					
TITLE	50 human secreted proteins					
JOURNAL	Patent: JP 2002539775-A 44 26-NOV-2002;					
COMMENT	HUMAN GENOME SCIENCES INC					
	OS Homo sapiens (human)					
	PN JP 2002539775-A/44					
	PD 26-NOV-2002					
	PF 09-WAR-2000 JP 2000606612					
	PR 19-WAR-1999 US 60/125360,11-JUN-1999 US 60/138626 PR					
	03-DEC-1999 US 60/168662					
	PI CRAIG A ROSEN,STEVEN M RUBEN,GEORGE KOMATSOUKIS PC					
	C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P1/00,A61P1/14, PC					
	A61P1/16,					
	PC A61P3/00,A61P3/06,A61P3/10,A61P7/00,A61P7/02,A61P7/06,A61P9/00, PC					
	A61P9/06,					
	PC A61P9/10,A61P9/10,A61P15/00,A61P17/02,A61P17/06,A61P19/00, PC					
	A61P25/00,					
	PC A61P25/14,A61P25/16,A61P25/24,A61P25/28,A61P27/00,A61P27/06,					
	PC A61P29/00,					
	PC A61P31/00,A61P31/04,A61P31/14,A61P31/16,A61P31/18,A61P31/20,					
	PC A61P31/22,					
	PC A61P33/00,A61P33/02,A61P33/06,A61P35/00,A61P35/02,A61P37/00,					
	PC A61P37/06,					
	PC A61P37/08,A61P43/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC					
	C12N1/21,					
	PC C12N5/10,C12P21/02,C12Q1/02,C12Q1/68,G01N33/68//C12P21/02, PC					
	C12R1/19),					

```

PC C12N15/00.A61K37/02.C12N5/00
CC 50 human secreted proteins
FT Key Location/Qualifiers
    1..1464
    /organism="Homo sapiens (human)"
FEATURES
    source
        1..1464
        Location/Qualifiers
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
ORIGIN
    Query Match          95.9%;   Score 75.8;   DB 6;   Length 1464;
    Best Local Similarity 97.5%;   Pred. No. 1.5e-11;
    Matches 77;   Conservative 0;   Mismatches 2;   Indels 0;   Gaps 0;
Qy 1 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
Db 1071 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 1130
Qy 61 AACAAACCGACTCTAAAAA 79
Db 1131 AACAAACCGACTCACAATA 1149
RESULT 2
AK223104 LOCUS 1757 bp mRNA linear PRI 26-APR-2005
DEFINITION Homo sapiens mRNA for secreted and transmembrane 1 precursor
            variant, clone: KAT04885.
ACCESSION AK223104
VERSION AK223104.1 GI:62897768
KEYWORDS FLI_CDNA; oligo capping.
SOURCE Homo sapiens
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE
    1
    Maruyama,K. and Sugano,S.
    Oligo-capping: a simple method to replace the cap structure of
    eukaryotic mRNAs with oligoribonucleotides
    Gene 138 (1-2), 171-174 (1994)
PUBMED 8125298
REFERENCE
    2
    Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and
    Sugano,S.
    Construction and characterization of a full length-enriched and a
    5'-end-enriched cDNA library
    Gene 200 (1-2), 149-156 (1997)
PUBMED 9373149
REFERENCE
    3 (bases 1 to 1757)
    Suzuki,Y., Sugano,S., Totoki,Y., Toyoda,A., Takeda,T., Sakaki,Y.,
    Tanaka,A. and Yokoyama,S.
    Direct Submission
    Submitted (22-APR-2005) Akiko Tanaka, RIKEN Yokohama Institute,
    Protein Research Group; 1-7-22 Suehiro, Tsurumi, Yokohama,
    Kanagawa, 230-0045, Japan (E-mail:aktanaka@riken.jp,
    URL:http://protein.gsc.riken.jp/, Tel:81-45-503-9452,
    Fax:81-45-503-9450)
    This work was supported in part by the National Project on Protein
    Structural and Functional Analysis, Ministry of Education, Culture,
    Sports, Science and Technology of Japan.
    Sumio Sugano, Yutaka Suzuki
    Laboratory of Functional Genomics Department of Medical Genome
    Sciences Graduate School of Frontier Sciences The University of
    Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639 Japan email:
    ssugano@k.u-tokyo.ac.jp
    URL:http://www.k.u-tokyo.ac.jp/index.html.en.
FEATURES
    source
        1..1757
        Location/Qualifiers
            /organism="Homo sapiens"
            /mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="KAT04885"
/cell_line="KAT III"
/clone_lib="KAT"
/notes="cloning vector: pME18SFL3
this clone is also named as hss001002732"
<123..869
/notes="Start codon is not identified."
/codon_start=1
/evidence="not experimental"
/product="secreted and transmembrane 1 precursor variant"
/protein_id="BAD96824.1"
/db_xref="GI:62897769"
/translation="MOTCLPAPGHVSQALGTLFLAASLSAQNEGWDSPICTEGVVS
VNGENTVNSCNISAFSHVNIKLRAHGESALFNEVAPGYSDRGWQLQVQGVQAL
VIKGRDSHAGLIMHVLGHQNNRQVTLVSGAEFQSPADTFGFWPFAVTVAFILL
VALVFWAYRCRRSQORREKFFLLEPQMKVAALRAGAQOGLSRASAEIWTDPSEPTF
RLPLAVFKPSPLGALELLSPQLFPYAADP"
CDS
    Query Match          95.9%;   Score 75.8;   DB 8;   Length 1757;
    Best Local Similarity 97.5%;   Pred. No. 1.5e-11;
    Matches 77;   Conservative 0;   Mismatches 2;   Indels 0;   Gaps 0;
Qy 1 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
Db 1650 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 1709
Qy 61 AACAAACCGACTCTAAAAA 79
Db 1710 AACAAACCGACTCACAATA 1728
RESULT 3
CQ923987 LOCUS 428 bp DNA linear PAT 23-NOV-2004
DEFINITION Sequence 5187 from Patent WO2004097052.
ACCESSION CQ923987
VERSION CQ923987.1 GI:56213928
KEYWORDS
    Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE
    1
    Burczynski,M.E., Twine,N.C., Slonim,D.K., Trepicchio,W.L.,
    Strahs,A., Immerman,F. and Dörner,A.J.
    Methods for prognosis and treatment of solid tumors
    Patent: WO 2004097052-A 5187 11-NOV-2004;
    Wyeth (US); Burczynski, Michael E. (US)
FEATURES
    source
        1..428
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
    misc_feature
        234..240
        /notes="n includes a, c, g, or t, or contains no
        nucleotide"
    misc_feature
        242..249
        /notes="n includes a, c, g, or t, or contains no
        nucleotide"
    misc_feature
        252..256
        /notes="n includes a, c, g, or t, or contains no
        nucleotide"
ORIGIN
    Query Match          94.7%;   Score 74.8;   DB 6;   Length 428;
    Best Local Similarity 97.4%;   Pred. No. 3.9e-11;
    Matches 76;   Conservative 0;   Mismatches 2;   Indels 0;   Gaps 0;
Qy 1 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60

```

Db 100 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 159

Qy 61 AACAAACCGACTCTAAAA 78
 |||||
 Db 160 AACAAACCGACTCAGAA 177
 |||||

RESULT 4
 LOCUS CS113124 522 bp DNA linear PAT 24-JUN-2005
 DEFINITION Sequence 142 from Patent WO2005054507.
 ACCESSION CS113124
 VERSION CS113124.1 GI:68224696
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Corfe B. and Chirakkal H.
 TITLE Gene screen
 JOURNAL Patent: WO 2005054507-A 142 16-JUN-2005;
 University of Sheffield (GB)
 LOCATION/Qualifiers
 FEATURES
 source 1..522
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 94.7%; Score 74.8; DB 6; Length 522;
 Best Local Similarity 97.4%; Pred. No. 3.8e-11;
 Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 Db 82 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 23
 |||||

Qy 61 AACAAACCGACTCTAAAA 78
 |||||
 Db 22 AACAAACCGACTCAGAA 5
 |||||

RESULT 5
 LOCUS AX780236 592 bp DNA linear PAT 14-JUL-2003
 DEFINITION Sequence 2393 from Patent WO03039443.
 ACCESSION AX780236
 VERSION AX780236.1 GI:32697230
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Dugas M., Ellis R., Broers B. and Mergenthaler S.
 TITLE Novel genetic markers for leukemias
 JOURNAL Patent: WO 03039443-A 2393 15-MAY-2003;
 Deutsches Krebsforschungszentrum (DE);
 Ludwig-Maximilians-Universitaet Muenchen (DE);
 PD Dr. Dr. (DE); Schoch, Claudia (DE); Kern, Wolfgang (DE)
 LOCATION/Qualifiers
 FEATURES
 source 1..592
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 94.7%; Score 74.8; DB 6; Length 592;
 Best Local Similarity 97.4%; Pred. No. 3.7e-11;

Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 Db 434 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 493
 |||||

Qy 61 AACAAACCGACTCTAAAA 78
 |||||
 Db 494 AACAAACCGACTCAGAA 511
 |||||

RESULT 6
 LOCUS BC017716 1761 bp mRNA linear PRI 29-JUN-2004
 DEFINITION Homo sapiens secreted and transmembrane 1, mRNA (cDNA clone
 MSC-21265 IMAGE:4413387), complete cds.
 ACCESSION BC017716
 VERSION BC017716.1 GI:17389345
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 1761)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Sapetson, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Vallalou, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 1761)
 Strausberg, R.
 Direct Submission
 Submitted (03-DEC-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapsa-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 cDNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdepaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 27 Row: h Column: 10
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 4506868.
 Location/Qualifiers

FEATURES

source	1. .1761 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:21265 IMAGE:4413387" /issue_types="Liver, adenocarcinoma" /clone_lib="NIH MGC_90" /lab_host="DH10B" /note="Vector: pCMV-SPORT6"	COMMENT	Protein Research Group; 1-7-22 Suehiro, Tsurumi, Yokohama, Kanagawa, 230-0045, Japan [E-mail:aktanaka@iken.jp. URL:http://protein.gsc.riken.jp/, Tel.81-45-503-9452, Fax:81-45-503-9450] This work was supported in part by the National Project on Protein Structural and Functional Analysis, Ministry of Education, Culture, Sports, Science and Technology of Japan. Sumio Sugano, Yutaka Suzuki Laboratory of Functional Genomics Department of Medical Genome Sciences Graduate School of Frontier Sciences The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639 Japan email: ssugano@k.u-tokyo.ac.jp URL: http://www.k.u-tokyo.ac.jp/index.html.en.
gene	1. .1761 /gene="SECTM1" /note="synonym: K12" /db_xref="GeneID:6398" /db_xref="MIM:602602" 93. .839 /gene="SECTM1" /codon_start=1 /product="secreted and transmembrane 1 precursor" /protein_id="AAH17716.1" /db_xref="GI:17389346" /db_xref="GeneID:6398" /db_xref="MIM:602602" /translation="MOTCPAPFGHVSQLGTLFLAASLSAQNEGWDSPICTEGVVS VSWGENTVMSCNISNAFHVNIKLRAHGOESAIFNEVAPGYFRDGMQLQVQGGVAQL VIKGARDSHAGLYMHVLGHORNROVTLVSGAEPQSDPTGFWPVPVAVTAVFILL VALVMPAWYRCRCQQORREKKFELLEPMKFAALRAGAQQGLSRASAEIWTDPSEPTP RPLALVFKPSPLGALELLSPQLFPYAADP"	FEATURES	Location/Qualifiers 1. .1766 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="KAT02983" /cell_line="KATO III" /clone_lib="KAT" /note="cloning vector: pME18SFL3 this clone is also named as hss001002587" <92. .838 /note="Start codon is not identified." /codon_start=1 /evidence="not_experimental" /product="secreted and transmembrane 1 precursor variant" /protein_id="BAD96788.1" /db_xref="GI:62897697"
CDS	94.7%; Score 74.8; DB 8; Length 1761; Best Local Similarity 97.4%; Pred. No. 2.9e-11; Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	CDS	<92. .838 /note="Start codon is not identified." /codon_start=1 /evidence="not_experimental" /product="secreted and transmembrane 1 precursor variant" /protein_id="BAD96788.1" /db_xref="GI:62897697"
ORIGIN	Query Match 94.7%; Score 74.8; DB 8; Length 1766; Best Local Similarity 97.4%; Pred. No. 2.9e-11; Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	ORIGIN	Query Match 94.7%; Score 74.8; DB 8; Length 1766; Best Local Similarity 97.4%; Pred. No. 2.9e-11; Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy	1 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60 	Qy	1 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db	1620 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1679 	Db	1619 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1678
Qy	61 AACAAACCGACTCTAAA 78 	Qy	61 AACAAACCGACTCTAAA 78
Db	1680 AACAAACCGACTCACAAA 1697	Db	1679 AACAAACCGACTCACAAA 1696
RESULT 7		RESULT 8	
LOCUS	AK223068 1766 bp mRNA linear PRI 26-APR-2005	LOCUS	CQ776411 2000 bp DNA linear PAT 11-MAR-2004
DEFINITION	Homo sapiens mRNA for secreted and transmembrane 1 precursor variant, clone: KAT02983.	DEFINITION	Sequence 97 from Patent EP1394274.
ACCESSION	AK223068	ACCESSION	CQ776411
VERSION	AK223068.1 GI:62897696	VERSION	CQ776411.1 GI:45379801
KEYWORDS	FLI CDNA; oligo capping.	KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens	SOURCE	Homo sapiens (human)
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 Maruyama,K. and Sugano,S. Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides Gene 138 (1-2), 171-174 (1994)	REFERENCE	1 Ohtani,N., Sugita,Y., Yamaya,M., Kubo,H., Negai,H. and Izuahara,K. Methods of testing for bronchial asthma or chronic obstructive pulmonary disease Patent: EP 1394274-A 97 03-MAR-2004; Genox Research, Inc. (JP) Location/Qualifiers 1. .2000 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"
AUTHORS	Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S.	AUTHORS	Ohtani,N., Sugita,Y., Yamaya,M., Kubo,H., Negai,H. and Izuahara,K.
TITLE	Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library	TITLE	Methods of testing for bronchial asthma or chronic obstructive pulmonary disease
JOURNAL	Gene 200 (1-2), 149-156 (1997)	JOURNAL	Patent: EP 1394274-A 97 03-MAR-2004;
PUBLISHED	9373149	PUBLISHED	Genox Research, Inc. (JP)
REFERENCE	3 (bases 1 to 1766)	REFERENCE	1. .2000 Location/Qualifiers 1. .2000 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"
AUTHORS	Suzuki,Y., Sugano,S., Totoki,Y., Toyoda,A., Takeda,T., Sakaki,Y., Tanaka,A. and Yokoyama,S.	AUTHORS	Ohtani,N., Sugita,Y., Yamaya,M., Kubo,H., Negai,H. and Izuahara,K.
TITLE	Direct Submission	TITLE	Methods of testing for bronchial asthma or chronic obstructive pulmonary disease
JOURNAL	Submitted (22-APR-2005) Akiko Tanaka, RIKEN Yokohama Institute,	JOURNAL	Patent: EP 1394274-A 97 03-MAR-2004;

```
ORIGIN
Query Match          94.7%; Score 74.8; DB 6; Length 2000;
Best Local Similarity 97.4%; Pred. No. 2.8e-11;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
    |||||
Db 1645 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCGACTCTAAAA 78
    |||||
Db 1705 AACAAACCGACTCAGAAA 1722

RESULT 9
LOCUS CQ861464 2000 bp DNA linear PAT 10-SEP-2004
DEFINITION Sequence 97 from Patent WO2004072265.
ACCESSION CQ861464
VERSION CQ861464.1 GI:51982453
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Burczynski, M., Twine, N., Dörner, A.J. and Trepicchio, W.L.
TITLE METHODS FOR MONITORING DRUG ACTIVITIES IN VIVO / I
JOURNAL Patent: WO 2004072265-A 97 26-AUG-2004;
Weth (US); Burczynski, Michael E. (US); Twine, Natalie C. (US);
Dörner, Andrew J. (US); Trepicchio, William L. (US)
FEATURES
source 1. .2000
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match          94.7%; Score 74.8; DB 6; Length 2000;
Best Local Similarity 97.4%; Pred. No. 2.8e-11;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
    |||||
Db 1645 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCGACTCTAAAA 78
    |||||
Db 1705 AACAAACCGACTCAGAAA 1722

RESULT 10
LOCUS CQ981630 2000 bp DNA linear PAT 25-JAN-2005
DEFINITION Sequence 485 from Patent EP1498424.
ACCESSION CQ981630
VERSION CQ981630.1 GI:58190920
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Rosenthal, A., Hermann, K., Heiden, E., Pilarsky, C., Brummendorf, T.,
Staub, E., Roepcke, S., Mennerich, D., Kinnemann, H. and Li, X.
TITLE Human nucleic acid sequences from lung tumours
JOURNAL Patent: EP 1498424-A 485 19-JAN-2005;
Hinzmann, Bernd (DE); Hermann, Klaus (DE); Heiden, Esmeralda (DE);
Rosenthal, Andre (DE)
FEATURES
source Location/Qualifiers
1. .2000
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match          94.7%; Score 74.8; DB 6; Length 2000;
Best Local Similarity 97.4%; Pred. No. 2.8e-11;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
    |||||
Db 1645 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCGACTCTAAAA 78
    |||||
Db 1705 AACAAACCGACTCAGAAA 1722

RESULT 11
LOCUS AR569106 2000 bp DNA linear PAT 14-DEC-2004
DEFINITION Sequence 3 from patent US 6762030.
ACCESSION AR569106
VERSION AR569106.1 GI:56569141
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2000)
AUTHORS Lyman, S.D. and Fanslow, W.C. III.
TITLE Ligand for CD7 and methods for use thereof
JOURNAL Patent: US 6762030-A 3 13-JUL-2004;
Immunex Corporation; Seattle, WA;
WOX;
FEATURES
source Location/Qualifiers
1. .2000
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match          94.7%; Score 74.8; DB 6; Length 2000;
Best Local Similarity 97.4%; Pred. No. 2.8e-11;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
    |||||
Db 1645 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCGACTCTAAAA 78
    |||||
Db 1705 AACAAACCGACTCAGAAA 1722

RESULT 12
LOCUS AX055560 2000 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 3 from Patent WO0073333.
ACCESSION AX055560
VERSION AX055560.1 GI:12228778
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Lyman, S.D. and Fanslow, W.C.
TITLE Ligand for cd7, and methods for use thereof
JOURNAL Patent: WO 0073333-A 3 07-DEC-2000;
Immunex Corporation (US)
FEATURES
source Location/Qualifiers
1. .2000
```



```
FEATURES
  source      Location/Qualifiers
              1..2180
              /organism="unknown"
              /mol_type="unassigned DNA"

ORIGIN
Query Match      94.7%; Score 74.8; DB 6; Length 2180;
Best Local Similarity 97.4%; Pred. No. 2.7e-11;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGGAATCCCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1841 CGGGAATCCCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1900
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 61 AACAAACCGACTCTAAAA 78
    ||||||||||||||||
Db 1901 AACAAACCGACTCAAAA 1918
    ||||||||||||
```

Search completed: December 7, 2005, 04:21:56
Job time : 703.666 secs

This Page Blank (uspr,

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 3355.79 Seconds
(without alignments)
4099.008 Million cell updates/sec

Title: US-09-980-046B-2
Perfect score: 294
Sequence: 1 cggccttaaggttcctctga.....aagagagtaattcccaaaa 294

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078125 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_ga1:*
10: gb_ga2:*
11: gb_ga3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	289.2	98.4	561	BF195984	7088e02.x
C 2	275.6	93.7	599	AW082845	AW082845 xc03h03.x
C 3	228.4	77.7	540	AQ167893	AQ167893 HS 2224 B
C 4	50	17.0	939	DU045370	DU045370 153083 To
C 5	49.6	16.9	807	CX948891	CX948891 UMC-bcl 0
C 6	49.4	16.8	909	CZ514707	CZ514707 GWM2-9TL2
C 7	49.4	16.8	928	CNS00DKY	AL071865 Drosophil
C 8	49.4	16.8	1101	CNS00EVL	AL069706 Drosophil
C 9	49.2	16.7	1101	CNS01450	AL103740 Drosophil
C 10	49	16.7	987	CNS014PQ	AL104456 Drosophil
C 11	48.2	16.4	897	CNS015FK	AL105386 Drosophil
C 12	48	16.3	979	CNS0161W	AL106190 Drosophil
C 13	48	16.3	1101	CNS0161L	AL108896 Drosophil
C 14	47.8	16.3	1101	CNS01219	AL101595 Drosophil
C 15	47.2	16.1	572	CX572895	CX572895 TTE000224
C 16	47	16.0	1432	AG381846	AG381846 Mus muscu
C 17	46.8	15.9	1201	CNS0107Y	AL098632 Drosophil
C 18	46.2	15.7	836	DN078399	DN078399 JGI_CASD1
C 19	46.2	15.7	836	CNS00K22	AL077933 Drosophil
C 20	46.2	15.7	964	BX426198	BX426198 BX426198
C 21	45.8	15.6	688	AG167069	AG167069 Pan trogl
C 22	45.8	15.6	811	AL514901	AL514901 AL514901

23	45.8	15.6	813	9	BZ512051	BZ512051 BOMQI3TF
C 24	45.6	15.5	479	8	DN344837	DN344837 LIB3551-0
C 25	45.6	15.5	652	10	CZ143061	CZ143061 OA BBA004
C 26	45.6	15.5	771	9	AV758683	AV758683 RP43-121h
27	45.4	15.4	1101	10	CNS00ZAL	AL097431 Drosophil
28	45.2	15.4	713	9	BZ043743	BZ043743 ljr30f06.
29	45.2	15.4	928	10	CNS00DKY	AL071865 Drosophil
30	45	15.3	850	10	CZ543128	CZ543128 SRAA-aad4
C 31	45	15.3	894	9	AQ621876	AQ621876 HS 3107 B
C 32	44.8	15.2	759	11	CNS06QXV	AL411257 T7 end of
33	44.6	15.2	285	3	BM893703	BM893703 i128f12.x
34	44.6	15.2	492	1	AU053286	AU053286 AU053286
C 35	44.4	15.1	481	11	DE137325	DE137325 Oryzias 1
36	44.4	15.1	744	7	CJ344067	CJ344067 CJ344067
37	44.4	15.1	951	11	CNS07BHQ	AL437892 T7 end of
C 38	44.4	15.1	992	10	CL144176	CL144176 ISB1-144B
39	44.4	15.1	1005	11	CNS07BO3	AL438121 T7 end of
40	44.4	15.1	1043	11	CNS07B49	AL437407 T3 end of
C 41	44.4	15.1	1298	10	AG360827	AG360827 Mus muscu
C 42	44.2	15.0	806	10	CNS008OW	AL052249 Drosophil
C 43	44.2	15.0	920	9	AZ671046	AZ671046 ETKKE23TR
C 44	44	15.0	624	8	CX573684	CX573684 TTR000312
45	44	15.0	747	10	AG364182	AG364182 Mus muscu

ALIGNMENTS

RESULT 1
BF195984/c

LOCUS
DEFINITION

BF195984 561 bp mRNA linear
7088e02.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:3643467 3',
similar to SW:STB1_BOVIN Q28208 SYNTAXIN BINDING PROTEIN 1 ; , mRNA
sequence.

ACCESSION
VERSION

BF195984
BF195984.1 GI:11083440

KEYWORDS
SOURCE

Homo sapiens
Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

1 (bases 1 to 561)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

TITLE

Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL, send email to:

info@image.llnl.gov

High quality sequence stop: 490.

Location/Qualifiers

1. 561

source

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3643467"

/lab_host="DH10B"

/clone_lib="NCI-CGAP Kid11"

/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with

a modified polylinker; Site 1: Not I; Site 2: Eco RI;

Plasmid DNA from the normalized library NCI-CGAP Kid3

prepared, and ss circles were made in vitro. Following HAP

purification, this DNA was used as tracer in a subtractive

hybridization reaction. The driver was PCR-amplified cDNAs

from a pool of 5,000 clones made from the same library

(cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo.

ORIGIN

Query Match 98.4%; Score 289.2; DB 2; Length 561;
Best Local Similarity 99.0%; Pred. No. 2e-53;
Matches 291; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGGCCTTAAGGTTCTCTGACAGTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 60
DB 294 CGGCCTTAAGGTTCTCTGACAGTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 235
QY 61 TCAAAATTAATCTTCCAGTGTCTTGTATCTCTCAATGTTATAGTGAATAATAATA 120
DB 234 TCAAAATTAATCTTCCAGTGTCTTGTATCTCTCAATGTTATAGTGAATAATAATA 175
QY 121 CATGCAGAAAAAGCATATAATCAACTATTGAATAACGAAATGTGAACCTCTATGTA 180
DB 174 CATGCAGAAAAAGCATATAATCAACTATTGAATAACGAAATGTGAACCTCTATGTA 115
QY 181 GCAACTCAGGCTGCTACCGTTTTTAATTTCTCTTAATCTTTTAAAAACAATACTATCT 240
DB 114 GCAACTCAGGCTGCTACCGTTTTTAATTTCTCTTAATCTTTTAAAAACAATACTATCT 55
QY 241 TTACTAAAGATGTTTCAGAGTAAGATAAATAACAGAGAGTAAATCCCAAAA 294
DB 54 TTACTAAAGATGTTTCAGAGTAAGATAAATAACAGAGAGTAAATCTCCAAA 1

RESULT 2

AW082845/c

LOCUS

DEFINITION

AW082845

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT: Robert Strausberg, Ph.D.

EMAIL: cgabbs-remail.nih.gov

TISSUE PROCUREMENT: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40UP from Gibco

High quality sequence stop: 417.

Location/Qualifiers

1. .599

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2583221"

/tissue_type="moderately differentiated adenocarcinoma"

/lab_host="DH10B"

/clone_lib="NCI CGAP Co21"

/notes="Organ: colon; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Normalized to Cot >500. Average insert size 1.04kb.

Normalized version of NCI_CGAP_Col8. Library constructed by Life Technologies."

ORIGIN

Query Match 93.7%; Score 275.6; DB 1; Length 599;
Best Local Similarity 98.3%; Pred. No. 2e-50;
Matches 289; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 CGGCCTTAAGGTTCTCTGACAGTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 60
DB 302 CGGCCTTAAGGTTCTCTGACAGTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 244
QY 61 TCAAAATTAATCTTCCAGTGTCTTGTATCTCTCAATGTTATAGTGAATAATAATA 120
DB 243 TCAAAATTAATCTTCCAGTGTCTTGTATCTCTCAATGTTATAGTGAATAATAATA 184
QY 121 CATGCAGAAAAAGCATATAATCAACTATTGAATAACGAAATGTGAACCTCTATGTA 180
DB 183 CATGCAGAAAAAGCATATAATCAACTATTGAATAACGAAATGTGAACCTCTATGTA 124
QY 181 GCAACTCAGGCTGCTACCGTTTTTAATTTCTCTTAATCTTTTAAAAACAATACTATCT 240
DB 123 GCAACTCAGGCTGCTACCGTTTTTAATTTCTCTTAATCTTTTAAAAACAATACTATCT 64
QY 241 TTACTAAAGATGTTTCAGAGTAAGATAAATAACAGAGAGTAAATCCCAAAA 294
DB 63 TTACTAAAGATGTTTCAGAGTAAGATAAATAACAGAGAGTAAATCTCCAAA 10

RESULT 3

AW0167893

LOCUS

DEFINITION

AW0167893

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

COMMENT

CONTACT: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence Tagged Connector

Plate: 2224 row: N column: 18

Class: BAC ends

High quality sequence stop: 540.

Location/Qualifiers

1. .540

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/clone="Plate=2224 Col=18 Row=N"

/sex="male"

/clone_lib="CIT Approved Human Genomic Sperm Library D"

/notes="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

ORIGIN

Query Match 77.7%; Score 228.4; DB 9; Length 540;
 Best Local Similarity 93.7%; Pred. No. 4.8e-40;
 Matches 238; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 CGGCCTTAAGTTCTCTGACAGTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 60
 Db 259 CGGCCTTAAGTTCTCTGACAGTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 318 ✓
 Qy 61 TCAAAATTATTTCTCCAGTGGTTTGTATCTCTCAATGTATATAGTGAATATAATA 120
 Db 319 TCAAAATTATTTCTCCAGTGGTCTTGTATCTCTCAATGTATATAGTGAATATAATA 378
 Qy 121 CATGCAGAAAAAGCATATAAACAATTTGTAATAACGAAATGTGAACCTCTATGTAAT 180
 Db 379 CATGGCACAAGCATATAAACAATTTGTAATAACGAAATGTGAACCTCTATGTAAT 438
 Qy 181 GCAACTCAGGCTCGCTACCGTTTTTAATTTCTCTCTAATCTTTTAAACAAATACTATCT 240
 Db 439 GCAACTCAGGCTCGCTACCGTTTTTAATTTCTCTCTAATCTTTTAAACAAATACTATCT 498
 Qy 241 TTACTTAAGATGTT 254
 Db 499 CTACTTAAGATGTT 512

RESULT 4
 DU045370/c
 LOCUS
 DEFINITION 153083 Tomato HindIII BAC Library Lycopersicon esculentum genomic clone LE_HBa0025020 5, genomic survey sequence.
 ACCESSION DU045370.1 GI:72471206
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Lycopersicon esculentum (Solanum lycopersicum)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 REFERENCE 1 (bases 1 to 939)
 AUTHORS Mueller, L.A., Buels, R.M., Wang, Y., Tankley, S.D., Giovannoni, J.J., Van Eck, J. and Stack, S.
 TITLE BAC end sequencing from three Solanum lycopersicon libraries
 JOURNAL Unpublished (2005)
 COMMENT Other GSSs: 153082
 Contact: Lukas Mueller
 Tanksley Lab, Dept. of Plant Breeding
 Cornell University
 251 Emerson Hall,
 Ithaca, NY 14853, USA
 Tel: 607-255-6557
 Fax: 607-255-6683
 Email: sgn-feedback@sgn.cornell.edu
 Insert Length: 104809 Std Error: 0.00
 Plate: 25 row: 0 column: 20
 Seq primer: T7
 Classes: BAC ends
 High quality sequence start: 51
 High quality sequence stop: 931.
 Location/Qualifiers
 1..939
 /organism="Lycopersicon esculentum"
 /mol_type="genomic DNA"
 /cultivar="Heinz 1706"
 /db_xref="taxon:4081"
 /clone="LE_HBa0025020"
 /lab_host="E. coli"
 /clone_lib="Tomato HindIII BAC Library"
 /note="Vector: pBelobAC11; Site_1: HindIII"

FEATURES

source

1..939
 /organism="Lycopersicon esculentum"
 /mol_type="genomic DNA"
 /cultivar="Heinz 1706"
 /db_xref="taxon:4081"
 /clone="LE_HBa0025020"
 /lab_host="E. coli"
 /clone_lib="Tomato HindIII BAC Library"
 /note="Vector: pBelobAC11; Site_1: HindIII"

ORIGIN

Query Match 17.0%; Score 50; DB 10; Length 939;
 Best Local Similarity 53.6%; Pred. No. 0.83;
 Matches 104; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 100 TTATAGTGAATATAATAATACATGCAGAAAAAGCATATAATCACTATTGAATAACGAA 159
 Db 848 TAAATTTAGGGAATATAGAAAAAATAATAATAATTAATCACTCCAAATTTGACAAATAAA 789
 Qy 160 AATGTGAACCTCTATGTAATCTGCAACTCAGGCTGCCTACCGTTTTTAATTTCTCTCTAAT 219
 Db 788 AATAGTAACCTCTAAAATCTACCTTTTATATATATTTTAAATTTTAAATTTTTCGATAAT 729
 Qy 220 CTTTAAAAACAAATACATCTCTTTACTAAAGATGTTTCAGAGTAGAATAAATAACAAAGAG 279
 Db 728 CTCAAAAATGAATAATTTGATTATAAAAAAATTTAAAAAATAATTAATAAAGATATA 669
 Qy 280 AGTAATTTCCCAAAA 293
 Db 668 TATAACCCCAAAA 655

RESULT 5

CX948891/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Ruminantia; Laurasiatheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2002)

Contact: DNA Core Facility (Bovine Project)

Animal Science - RS Prather

University of Missouri-Columbia

M16 Medical Sciences Bldg., Columbia, MO 65212, USA

Tel: (573)882-0428

Fax: (573)884-5552

Email: bovine@net.missouri.edu

POLYA=Yes.

Location/Qualifiers

1..807

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/clone_lib="bcl"

/note="Funding: The production of ESTs submitted in this project was funded by USDA Grant MRI-2002-03476 entitled 'Bovine ESTs: Focus on Female Reproduction' to RS Prather (Primary Investigator), E Antoniou, HA Garverick, JA Green, MC Lucy, RM Roberts, MF Smith and RS Youngquist. Genetic Source: Heifers for the project were purchased from Circle A Ranch, Iberia, MO (http://www.circlea ranch.com/home.html). These heifers, while not registered have known Angus pedigrees going back at least 4 generations. Samples collected: The samples consisted of the following: germinal vesicle-stage oocytes; in vitro derived embryos (2-cell, morula, blastocyst and nuclear transfer blastocyst); in vivo blastocysts and conceptuses (days 8, 14, 16 and 18); corpora lutea (days 3, 5, 8, 14, 16, 18 and 35); ovarian follicles (days 0, non-recruited, recruited, early selected and preovulatory); oviduct (days 0, 3 and 5); endometrium (days 5, 8, 14, 16, 18 and 35); and placenta/embryo from day 35 conceptuses. Expanded descriptions of how the tissues were collected can be found at the following URL: http://genome.rnet.missouri.edu/Bovine/Methods.html.

FEATURES

source

Library construction (Standard Protocol): All procedures have been described in detail elsewhere (Soares et al., 1994; Bonaldo et al., 1996; Jiang et al., 2001). Total cellular RNA from each sample was isolated by using STAR-60 reagent (Tel-Test, Friendswood, TX) and the poly(A)+ RNA was obtained by two rounds of purification with the Oligotex mRNA isolation kit (Qiagen) according to the manufacturer's instructions. The libraries were constructed essentially as described by the manufacturer's instructions provided with the SuperScript Plasmid System (Invitrogen, cat. no. 18248-013). Briefly, 1mcg of poly(A)+ RNA was annealed at c37 degrees with 10mcg of NotI-tag-dT18 oligonucleotide (GCTGCTCGGCGC-tag-T18) and reverse transcribed at c37 degrees with SuperScript II (Invitrogen) reverse transcriptase (Jiang et al., 2001). The 'tag' represents a tissue/stage-specific ten-base sequence identifier (<http://genome.illowa.edu/pubsoft/software.html>) present in the oligonucleotide used to prime first-strand synthesis. Second strand synthesis was performed with T4 DNA polymerase in the presence of DNA ligase and RNase H. After second strand synthesis, the double-stranded cDNAs were ligated to SalI adapters (Invitrogen-Life Technologies) and digested with NotI. The cDNAs were size selected by passage through cDNA size fractionation columns (Invitrogen-Life technologies). The cDNAs derived from each developmental stage of a particular tissue were mixed on an equimolar basis and ligated directionally into the NotI and SalI sites of the pCMV-SPORT6 vector (Invitrogen). After ligation of the inserts, the plasmids were electroporated into DH10B bacteria. Library construction (PCR Protocol): The amount of mRNA that was recovered from oocytes and embryos was quite limiting and was not sufficient for library production with the standard protocol. Therefore, PCR-based protocol was utilized for producing libraries from sources in which the amount of extracted mRNA was small (oocytes and embryos). Poly-A RNA was isolated by using the MicroPoly(A) Pure kit from Ambion (cat. # 1918). The mRNA was reverse transcribed with a NotI-tag-dT18 oligonucleotide and a SMART oligonucleotide (Clontech) modified to contain a SalI site to generate full-length cDNA with a sequence complementary to the SMART oligonucleotide. Sequences within the SMART and dT oligonucleotides were used as primers to amplify the cDNAs by PCR with pfu turbo polymerase (Stratagene). The resulting PCR products were purified, digested with NotI and SalI and size fractionated by using Chroma Spin-1000 columns (Clontech). Purified cDNA from each PCR reaction was quantitated and mixed on an equimolar basis for ligation into the pCMV-SPORT6 vector. Preliminary Library Characterization: Randomly chosen clones from each library were analyzed by restriction digestion to determine average insert size (96 clones) and by sequencing (-4 96-well plates) to confirm library quality [e.g. the presence of short polyA+ tails, genomic DNA contamination (must be <1%), ribosomal RNA clones (must be <1%), etc.] and to provide a sequence database representing the predominant clones in each library. The clones were sequenced at the University of Missouri-Columbia DNA Core Facility. After production of the libraries, equal numbers of recombinants from each library were pooled to produce a single mixed library (mega-library) for more extensive sequencing. Bioinformatics work was performed by GK Springer's bioinformatics group in Computer Science at the University of Missouri-Columbia. Clone Requests: Requests for clones should be made to the Director of the University of Missouri DNA Core facility at: bovine@net.missouri.edu. Bonaldo MF, Lennon G, Soares MB.. Normalization and Subtraction: Two approaches to facilitate gene discovery. Genome Res, 1996; 6:791-806. Jiang H, Bivens NJ, Ries JE, Whitworth KM, Green JA, Forrester LJ, Springer GK, Diddion BA, Mathialagan N, Prather RS, Lucy MC (2001) Constructing

cDNA libraries with fewer clones that contain long poly(dA) tails. Biotechniques 31:38-42. Soares MB, MF Bonaldo, P Jelene, L Su, L Lawton, A Efstratiadis. 1994. Construction and characterization of a normalized cDNA library. Proc Natl Acad Sci, 91:9228-9232. TAG ISSUE=Day 14 CL +8h prostaglandin F2-alpha TAG_SEQ=GAAGTGCTCC"

ORIGIN

Query Match 16.9%; Score 49.6; DB 8; Length 807;
Best Local Similarity 50.9%; Pred. No. 1;
Matches 118; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 56 TTACTTCAATATTCTTCAGATGGTTTGGTTATCTTCTCAATGTTATAGTGAATATA 115
DB 421 TTTCTCGATAAAATTTTTTTTTTGGCAATTTTAAAGCTGTGCATTTGTTACTTTCAA 362

QY 116 TAATACATGCAGAAAAAGCATAAATCAACTATTGTAATAACGAAAAATGTGAACCTCTATG 175
DB 361 TTTTAAGTCATGAAATATATATGCAATTTATATGTAATAAGTTACATGAGAAATTTCCCTG 302

QY 176 TAACTGCAACTCAGGCTGCCTACCGTTTAAATTTCTCTTAATCTTTAAAAACAATAC 235
DB 301 TAACTACTGTCGAGAAATTTAAGATGTTTATATATCTTTAAAAATATTTTTCAGAGCTAACTTA 242

QY 236 TATCTTTTAAAGATGTTTTCAGAGTAAGTAAATACAAAGAGAGTAAATTC 287
DB 241 TTTATATGTTTTTACAGTTTTCAGAGAAAAGTACAAGCAAGTTGTAAATTC 190

RESULT 6

CZ514707 909 bp DNA linear GSS 11-MAY-2005
LOCUS GMW2-91L21a.g1 GMW2 Glycine max genomic, genomic survey sequence.
DEFINITION
ACCESSION CZ514707
VERSION CZ514707.1 GI:63239622
KEYWORDS GSS.
SOURCE Glycine max (soybean)
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 909)
AUTHORS Nunberg,A., Bedell,J.A., Citek,R.W., Robbins,D., McMenamy,J., Peterson,S., Jones,J., Fries,J., Budiman,M.A., Nguyen,H., Clifton,S.W. and Stacey,G.
TITLE BAC end sequences from Glycine max Williams 82 cultivar genome
JOURNAL Unpublished (2005)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Class: BAC ends.

FEATURES
source 1..909
/organism="Glycine max"
/mol_type="genomic DNA"
/db_xref="taxon:3847"
/clone_lib="GMW2"
/notes="Glycine max BAC library"

ORIGIN

Query Match 16.8%; Score 49.4; DB 10; Length 909;
Best Local Similarity 52.7%; Pred. No. 1.1;
Matches 107; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 35 AGATCAGGATTGAGCATCTTACTTCAATTTCTTCCAGTGGTTTGTTATCTTCT 94
DB 50 AAATAAGAAGTGTGCTTCTTCACTTTTAAAAAACCCCTTTAGTCTGCTTCTTCTTTT 109
QY 95 CAATGTTTAGTGAATATAATAATACATGCAGAAAAAGCATAAATCACTTGAATA 154

```

Db      110 GAGTGGGTGGTGTGTTTGTGAGAAGAGATAAAAAACAACAGAAATAGATAATTAATA 169
Qy      155 ACAGAAATGTCGAACCTCTATGCTCAACTCAGGCTGCCTACCGTGTCTTAAATTTCTCT 214
Db      170 ATGTAATTTATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTCT 229
Qy      215 CTAATCTTTAAAAACAATACTA 237
Db      230 CAACCTTTTAACTATAAATA 252

RESULT 7
CNS00DKY/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR27A24 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL071865
VERSION
AL071865.1 GI:4948170
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 928)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
Location/Qualifiers
1..928
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR27A24"
/clone_lib="RPCI-98"
/note="end : T7"

ORIGIN
Query Match 16.8%; Score 49.4; DB 10; Length 928;
Best Local Similarity 31.8%; Pred. No. 1.1;
Matches 92; Conservative 73; Mismatches 122; Indels 2; Gaps 1;

Qy      8 AAGGTTCCTCTGACAGTGTCTCAAGAAGATCAGGATTTGAGCATCTCTACTTCAAAAT 67
Db      862 AAATTTCAAAAMAMAMMMCMVMMMMMMMMMMMMMMMMMMMMMMHHHHHHHHHTTTTTW 803
Qy      68 ATTCTCCAGTGGTTTGTATCTTCTCAATGTTATAGTGAATAATAATAATACATGAG 127
Db      802 TTTTCTTTTCTTTTMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMNNNNNNCHWAWATMMW 743
Qy      128 AAAAAAGCAATAATCAACTATTGAATACGAAAAATGTGAACCTCTATGTAAGTCACTC 187
Db      742 WWATATTTTATTTWAWATAAAAAAATAAAATTAATTTTWTWTATTTTAA 683
Qy      188 AGGCTGCCTACCGTTT- TTTAATTTCTCTCTAATCTTTAAAAACAATACTACTTCTTACT 245

```

```

Db      682 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTATTTAT 623
Qy      246 AAAGATGTTTTCAGAGTAAGATAAAATACAAAGAGAGTAATTTCCCAAAA 294
Db      622 ATTTTWTATTATTTATTTATTTATTTATTTTATTTTATTTTATTTATTTATTTAATA 574

RESULT 8
CNS00EVL
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC:
BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL069706
VERSION
AL069706.1 GI:4949849
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR29B23"
/clone_lib="RPCI-98"
/note="end : T7"

ORIGIN
Query Match 16.8%; Score 49.4; DB 10; Length 1101;
Best Local Similarity 37.9%; Pred. No. 1.1;
Matches 92; Conservative 40; Mismatches 111; Indels 0; Gaps 0;

Qy      52 TCTCTTACTTCAAAATTTCTTCCAGTGGTGTTCCTTATCTTCTCAATGTTATAGTGAAT 111
Db      633 TTTAATAATTTAATTTATTTATTTAATAATTTAATTTAATAATTTAATAATAA 692
Qy      112 ATAATAATACATGCAGAAAAAAGCATAAATCAACTATTGAAATAACGAAAAATGTGAACCTC 171
Db      693 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 752
Qy      172 TATGTAAGTCAACTCAGGCTGCCTACCGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 231
Db      753 AATATWATAATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 812
Qy      232 ATACTATCTTTACTTAAGATGTTTTCAGATGTAAGATAAATAACAAGAGAGTAATTTCCCAA 291
Db      813 ATAWATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 872

```

```
QY 292 AAA 294
Db 873 WAA 875

RESULT 9
CNS0145U/c
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
BACN1016 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL103740
VERSION AL103740.1 GI:5615351
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
REFERENCE
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segret@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.

FEATURES
source
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN1016"
/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/note="end : T7"

ORIGIN
Query Match 16.7%; Score 49.2; DB 10; Length 1101;
Best Local Similarity 39.6%; Pred. No. 1.2;
Matches 93; Conservative 11; Mismatches 131; Indels 0; Gaps 0;

QY 60 TTCAAAATTATTCTCCAGTGGTTTGTGTTATCTCTCAATGTTATAGTGAATAATAAT 119
Db 368 TTTNNNNNTTTTNTNNNNNTNNNNNTNNNNNTNNNNNTNNNNNTNNNNNTNNNNNT 309

QY 120 ACATGCGAGAAAAAGCATAAATCAACTATTGAATAACGAAAAATGTGAACCTCTATGTAAC 179
Db 308 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 249

QY 180 TGCACCTCAGGCTCGCTACCGCTTTTAAATTTCTCTCTAACTTTAAAAACAATACATC 239
Db 248 WTAATTATAATTTNNNNNTNNNTNNNNNTNNNNNTNNNNNTNNNNNTNNNNNTNNNNNT 189

QY 240 TTTACTAAGATGTTTCAGAGTAAGATAAATAACAGAGAGTAAATCCCAAAA 294
Db 188 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 134

RESULT 10
CNS014PQ/c
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN12P22 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL104456
VERSION AL104456.1 GI:5616067
```

```
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 987)
REFERENCE
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segret@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.

FEATURES
source
Location/Qualifiers
1..987
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN12P22"
/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/note="end : SP6"

ORIGIN
Query Match 16.7%; Score 49; DB 10; Length 987;
Best Local Similarity 32.4%; Pred. No. 1.4;
Matches 72; Conservative 55; Mismatches 95; Indels 0; Gaps 0;

QY 64 AATTATTCTCCAGTGGTTTGTGTTATCTCTCAATGTTATAGTGAATAATAATAATACAT 123
Db 913 WWWWWHTTTHWWTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 854

QY 124 GCAGAAAAAGCATAAATCAACTATTGAATAACGAAAAATGTGAACCTCTATGTAACGCA 183
Db 853 TTTATATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 794

QY 184 ACTCAGGCTGCCTACCGCTTTTAAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 243
Db 793 TAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 734

QY 244 CTAAAGATGTTTCAGAGTAAGATAAATAACAAAGAGAGTAAT 285
Db 733 TTTWTWTWTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 692

RESULT 11
CNS015FK
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN13G14 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL105386
VERSION AL105386.1 GI:5617400
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 897)
REFERENCE
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segret@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 128.321 Seconds
(without alignments)
4072.612 Million cell updates/sec

Title: US-09-980-046B-2

Perfect score: 294

Sequence: 1 cggccttaaggcttcctctga.....aagagagtaattcccaaaaa 294

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA.*
- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
 - 2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
 - 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
 - 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
 - 5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
 - 6: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
 - 7: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
 - 8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
 - 9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	288.2	98.0	84425	3	US-09-949-016-17402
2	43.6	14.8	65415	3	US-09-949-016-16699
3	43	14.6	1684976	3	US-08-916-421B-1
4	43	14.6	1684976	3	US-09-692-570-1
5	42.4	14.4	601	3	US-09-949-016-198281
6	42.4	14.4	767677	3	US-09-949-016-12147
7	42.4	14.4	767677	3	US-09-949-016-17361
8	42	14.3	601	3	US-09-949-016-198280
9	40.8	13.9	613	2	US-08-630-822A-98
10	40.8	13.9	613	2	US-09-005-089-98
11	40.8	13.9	613	3	US-09-171-156A-47
12	40.8	13.9	613	3	US-09-004-730A-47
13	40.8	13.9	613	3	US-08-981-799A-47
14	40.6	13.8	832	3	US-09-621-976-2813
15	40.4	13.7	50000	3	US-09-662-254B-23
16	40.2	13.7	601	3	US-09-949-016-201502
17	40.2	13.7	601	3	US-09-949-016-201726
18	40.2	13.7	2680	3	US-09-949-016-5657
19	40.2	13.7	3349	3	US-09-949-016-17399
20	40.2	13.7	52174	3	US-09-949-002-573
21	40.2	13.7	52289	3	US-09-949-002-720
22	40.2	13.7	86639	3	US-09-949-016-17397
23	39.8	13.5	601	3	US-09-949-016-201501
24	39.8	13.5	601	3	US-09-949-016-201727

c	25	39.4	13.4	32392	3	US-09-662-254B-27	Sequence 27, Appl
c	26	39.4	13.4	50000	3	US-09-662-254B-23	Sequence 23, Appl
c	27	39	13.3	601	3	US-09-949-016-185650	Sequence 185650,
c	28	39	13.3	1052	3	US-09-976-594-819	Sequence 819, App
c	29	39	13.3	168104	3	US-09-949-016-12026	Sequence 12026, A
c	30	39	13.3	168105	3	US-09-949-016-16554	Sequence 16554, A
c	31	38.8	13.2	601	3	US-09-949-016-170445	Sequence 170445, A
c	32	38.4	13.1	97376	3	US-09-949-016-16093	Sequence 16093, A
c	33	38.4	13.1	387902	3	US-09-949-016-14543	Sequence 14543, A
c	34	38.4	13.1	421883	3	US-09-949-016-12557	Sequence 12557, A
c	35	38	12.9	314798	3	US-09-949-016-13539	Sequence 13539, A
c	36	37.8	12.9	16573	3	US-09-949-016-14876	Sequence 14876, A
c	37	37.8	12.9	29133	3	US-09-949-016-12694	Sequence 12694, A
c	38	37.8	12.9	38479	3	US-09-949-016-16730	Sequence 16730, A
c	39	37.6	12.8	3942	3	US-09-601-198-50	Sequence 50, Appl
c	40	37.6	12.8	5448	3	US-09-620-312D-246	Sequence 246, App
c	41	37.6	12.8	5532	3	US-09-620-312D-245	Sequence 245, App
c	42	37.4	12.7	601	3	US-09-949-016-185649	Sequence 185649,
c	43	37.4	12.7	2877	3	US-09-235-103-1	Sequence 1, Appli
c	44	37.4	12.7	113966	3	US-09-949-016-12277	Sequence 12277, A
c	45	37.4	12.7	113967	3	US-09-949-016-17051	Sequence 17051, A

ALIGNMENTS

RESULT 1

US-09-949-016-17402

; Sequence 17402, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: Fast-Seq for Windows Version 4.0

; SEQ ID NO 17402

; LENGTH: 84425

; TYPE: DNA

; ORGANISM: Human

; US-09-949-016-17402

Query Match 98.0% Score 288.2; DB 3; Length 84425;

Best Local Similarity 99.0%; Pred. No. 3.8e-61;

Matches 290; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGGCTTAAGGTTCTCTGACAGTTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 60

Db 53140 CGGCTTAAGGTTCTCTGACAGTTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 53199

Qy 61 TCAAAATTAATTTCTCCAGTGGTTTGTATCTTCTCAATGTTATAGTGAATATAATA 120

Db 53200 TCAAAATTAATTTCTCCAGTGGTTTGTATCTTCTCAATGTTATAGTGAATATAATA 53259

Qy 121 CATGCAAAAAGCATATAATCAACTATTCGAATACCAAAATGGAACCTCTATGTA 180

Db 53260 CATGCAAAAAGCATATAATCAACTATTCGAATACCAAAATGGAACCTCTATGTA 53319

Qy 181 GCAACTCAGGCTGCTACCGTTTTTAATTTCTCTCTATCTTTAAAAACAATACTATCT 240

Db 53320 GCAACTCAGGCTGCTACCGTTTTTAATTTCTCTCTATCTTTAAAAACAATACTATCT 53379

Qy 241 TTACTAAAGATGTTTTCAGAGTAGAATAAAATACAAAGAGAGTAATTCCTCAAAA 293

Db 241 TTACTAAAGATGTTTTCAGAGTAGAATAAAATACAAAGAGAGTAATTCCTCAAAA 293

```
Db 53380 TTACTAAGATGTTTCAGAGTAAGATAAAATACAAAGAGAGTAATTCCTCAA 53432

RESULT 2
US-09-949-016-16699/c
; Sequence 16699, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16699
; LENGTH: 65415
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16699

Query Match 14.8%; Score 43.6; DB 3; Length 65415;
Best Local Similarity 48.8%; Pred. No. 0.27;
Matches 118; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 2 GGCCTTAAGTTCCTCGACAGTGTCTCAAGAGATCAGGATTTGAGCATCTCTTACTT 61
Db 39198 GGCATCAAGTCCTCTCGATGGTTTCTTTTCAGTTCCTTATTTAGATCTTCCTT 39139
QY 62 CAAATATTCTTCAGTGGTGTGTTGTTATCTCTCAATGTTATAGTGAATATAATATAC 121
Db 39138 CAGACTCGAATTAAGTCTGGTGTGTTAGTCAAGATGCAATTTGATCTTTAAAAATGGA 39079
QY 122 ATGCAGAAAAAGCATAAATCAACTATTGTAATACGAAAAATGGAACCTCTATGTAACGT 181
Db 39078 GGAAGAAGAAGAAATTAATATATGACACATTTTTCAAAGTTTGAAGTCCAATGTAACCT 39019
QY 182 CAACTAGGCTGCTACCGTTTAAATTCCTCTAATCTTTAAACAAATATATCTT 241
Db 39018 TATGTTACTGCTTTAATAACTTTTTTTTTTGAATACTCTAAATAATACCCATATACCT 38959
QY 242 TA 243
Db 38958 CA 38957

RESULT 3
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Buit et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; jannaschii
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PR275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
```

```
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
```

```

; LOCATTN: (1602912)..(1602912)
; NAME/KEY: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATTN: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATTN: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATTN: (1664854)..(1664855)
; OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Query Match      14.6%; Score 43; DB 3; Length 1664976;
Best Local Similarity 52.5%; Pred.No. 0.68;
Matches 94; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy      84  TGTATCTCTCCTCAATGTTATAGTGAATATAATAATCATGACAGAAAAAGCATAAATCA 143
Db      411276  TGTCTTTTCATAAAGTAGGATATAATTCATTTAAATCTCCATCTATATATCTCAATAA 411335

Qy      144  ACTATTGAATAAACGAAAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCTACCGTTT 203
Db      411336  ACTCTCATTTAAAAATAAGCTATAGCCCTCATTAACCTAATAGCTCTTAAACGGATT 411395

Qy      204  TTAATTTCTCTCTAAATCTTTAAAAACAAATACTATCTTTACTAAAGATGTTTTCAGAGTA 262
Db      411396  TTTTACCCTTACTTTAACTTTAAATTTTATACCTTCTTTCTTTAAATGTTTATAAAGTAA 411454

RESULT 4
US-09-692-570-1
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Meth
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:

```

```
;/ NAME/KEY: misc feature
;/ LOCATION: (98120)..(98120)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (98159)..(98159)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (98239)..(98239)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (98266)..(98266)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (98343)..(98343)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (103998)..(103998)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (148948)..(148948)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (163385)..(163385)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (191989)..(191989)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (191995)..(191995)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (231980)..(231980)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (234187)..(234187)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (234220)..(234220)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (234814)..(234814)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (309398)..(309398)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (309418)..(309418)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (312837)..(312837)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (312993)..(312993)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
```

```
;/ LOCATION: (319226)..(319226)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (559167)..(559167)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (559241)..(559241)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (600992)..(600992)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (622708)..(622708)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (657081)..(657081)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (657203)..(657203)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (674435)..(674435)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (682442)..(682442)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (713652)..(713652)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (741684)..(741684)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (779455)..(779455)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (779676)..(779676)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (855539)..(855539)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (871619)..(871619)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (1084830)..(1084830)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (1096846)..(1096846)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (1119881)..(1119881)
;/ OTHER INFORMATION: n equals a, t, c, or g
;/ FEATURE:
;/ NAME/KEY: misc feature
;/ LOCATION: (1130881)..(1130881)
```

OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g

Query Match 14.4%; Score 43; DB 3; Length 1664976;
Best Local Similarity 52.5%; Pred. No. 0.68;
Matches 94; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
Qy 84 TGTTCCTCTCTCAATGTTATAGTGAATATAATATACATGCAGAAAAAGCATAAATCA 143
Db 411276 TGTCTTTTCATAAAGTAGGATAAATTACATTAATAATCTCCATCTATATATCTCAATAA 411335
Qy 144 ACTATTGATTAACGAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCTACCGTTT 203
Db 411336 ACTCTCATTTAAATAAAGCTATAGCCCTCAATAACTAACTCTATAGCTCTAAACGGATT 411395
Qy 204 TTAATTTCTCTCTAATCTTTTAAACAAATACTATCTTTACTAAAGATGTTTCAGAGTA 262
Db 411396 TTTTACCTTACTTAACTTTTAAATTTTATACCTTCTTTCTTTAAATAGTTTATAACTA 411454

RESULT 5
US-09-949-016-198281/c
; Sequence 198281, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 198281
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-198281

Query Match 14.4%; Score 42.4; DB 3; Length 601;
Best Local Similarity 50.0%; Pred. No. 0.23;
Matches 106; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
Qy 44 TTTGAGCATCTCTTACTTCAAAATATTCTCCAGTGGTTTTGTTATCTTCTCAATGTTAT 103
Db 315 TTTAAGTAACATCATGCTTAACTATATTTTCATATGTCAAAGGGATATGTTTGAAGCTAG 256
Qy 104 AGTGAATATAATAATACATGCAGAAAAAGCATAAATCAACTATGTAATCAACGAAATG 163
Db 255 AGTCAAAATCTTTCTCATGTAAGAATGTGTAATGATTATATTAATAAAAAAATG 196
Qy 164 TGAACCTCTATGTAACCTGCAACTCAGGCTGCTACCGTTTTTAATTTCTCTCAATCTTT 223
Db 195 TTGTAGAACAAATATTTGCAAAATACATTTGTTTCATATGTTTCAAAATAAGTTATAAAAAAGT 136
Qy 224 AAAAAACAAATACTATCTTTTACTAAAGATGTTT 255

Db 135 AAACAGAAATATAGTTTTTTTCCCAACTTCCTTT 104

RESULT 6

US-09-949-016-12147
; Sequence 12147, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12147
; LENGTH: 767677
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(767677)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12147

Query Match 14.4%; Score 42.4; DB 3; Length 767677;
Best Local Similarity 50.0%; Pred. No. 0.83; Indels 0; Gaps 0;
Matches 106; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Qy 44 TTTGAGCATCTCTTACTTCAAAATATTCTCCAGTGGTTTTGTTATCTTCTCAATGTTAT 103
Db 668111 TTTAAGTAACATCATGCTTAACTATATTTTCATATGTCAAAGGGATATGTTTGAAGCTAG 668170
Qy 104 AGTGAATATAATAATACATGCAGAAAAAGCATAAATCAACTATGTAATCAACGAAATG 163
Db 668171 AGTCAAAATCTTTCTCATGTAAGAATGTGTAATGATTATATTAATAAAAAAATG 668230
Qy 164 TGAACCTCTATGTAACCTGCAACTCAGGCTGCTACCGTTTTTAATTTCTCTCAATCTTT 223
Db 668231 TTGTAGAACAAATATTTGCAAAATACATTTGTTTCATATGTTTCAAAATAAGTTATAAAAAAGT 668290
Qy 224 AAAAAACAAATACTATCTTTTACTAAAGATGTTT 255
Db 668291 AAACAGAAATATAGTTTTTTTCCCAACTTCCTTT 668322

RESULT 7

US-09-949-016-17361
; Sequence 17361, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
US-09-949-016-17361

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 34.4809 Seconds
(without alignments)
4072.612 Million cell updates/sec

Title: US-09-980-046b-3

Perfect score: 79

Sequence: 1 cgggaatcccccttcagttc.....aaacaacgcgactctaaaaa 79

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/PPUS COMB.seq.*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
- 9: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	74.8	94.7	601	3	US-09-949-016-26334
C 2	74.8	94.7	601	3	US-09-949-016-105203
C 3	74.8	94.7	2000	3	US-09-997-165-3
4	74.8	94.7	2000	3	US-09-949-016-426
5	74.8	94.7	2001	3	US-09-949-016-2936
6	74.8	94.7	2180	2	US-08-755-559-2
7	74.8	94.7	2180	3	US-09-210-474-2
8	74.8	94.7	2180	3	US-09-539-774-2
9	74.8	94.7	16738	3	US-09-949-016-12168
10	74.8	94.7	16738	3	US-09-949-016-14678
11	28.2	35.7	45983	3	US-09-949-016-17010
12	27.8	35.2	818128	3	US-09-949-016-14546
13	27.8	35.2	818128	3	US-09-949-016-14547
14	27.8	35.2	818128	3	US-09-949-016-14548
15	27.8	35.2	818128	3	US-09-949-016-14549
16	27.8	35.2	818128	3	US-09-949-016-14550
17	27.8	35.2	818128	3	US-09-949-016-14551
18	27.8	35.2	818128	3	US-09-949-016-14552
19	27.8	35.2	818128	3	US-09-949-016-14553
20	27.8	35.2	818128	3	US-09-949-016-14554
21	27.8	35.2	818128	3	US-09-949-016-14555
22	27.8	35.2	818128	3	US-09-949-016-14556
23	27.8	35.2	818128	3	US-09-949-016-14557
24	27.8	35.2	818128	3	US-09-949-016-14558

ALIGNMENTS

RESULT 1

US-09-949-016-26334/c
; Sequence 26334, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26334
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-26334

Query Match 94.7%; Score 74.8; DB 3; Length 601;
Best Local Similarity 97.4%; Pred. No. 3e-15; 2; Indels 0; Gaps 0;
Matches 76; Conservative 0; Mismatches 2;

Qy 1 CGGGAATCCCCCTTCAGTTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 129 CGGGAATCCCCCTTCAGTTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 70

Qy 61 AACAAACCGACTCTAAAA 78
Db 69 AACAAACCGACTCAGAAA 52

RESULT 2

US-09-949-016-105203/c
; Sequence 105203, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307

Sequence 14559, A
Sequence 14560, A
Sequence 14561, A
Sequence 14562, A
Sequence 14564, A
Sequence 14565, A
Sequence 14566, A
Sequence 14567, A
Sequence 16148, A
Sequence 666, App
Sequence 8073, App
Sequence 17314, A
Sequence 18604, A
Sequence 34586, A
Sequence 142500,
Sequence 15891, A
Sequence 12372, A
Sequence 15772, A
Sequence 18104, A
Sequence 50, Appl
Sequence 8100, Ap

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105203
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-105203

Query Match          94.7%; Score 74.8; DB 3; Length 601;
Best Local Similarity 97.4%; Pred. No. 3e-15;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCGCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
DB 129 CGGGAATCCCGCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 70
QY 61 AACAAACCGACTCTAAAA 78
DB 69 AACAAACCGACTCAAAA 52

RESULT 3
US-09-997-165-3
; Sequence 3, Application US/09997165
; Patent No. 6762030
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Fanslow, William C.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)..(865)
US-09-997-165-3

Query Match          94.7%; Score 74.8; DB 3; Length 2000;
Best Local Similarity 97.4%; Pred. No. 4e-15;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCGCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
DB 1645 CGGGAATCCCGCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704
QY 61 AACAAACCGACTCTAAAA 78
DB 1705 AACAAACCGACTCAAAA 1722

RESULT 4
US-09-949-016-426
; Sequence 426, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 426
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-426

Query Match          94.7%; Score 74.8; DB 3; Length 2000;
Best Local Similarity 97.4%; Pred. No. 4e-15;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCGCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
DB 1645 CGGGAATCCCGCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704
QY 61 AACAAACCGACTCTAAAA 78
DB 1706 AACAAACCGACTCAAAA 1723

RESULT 5
US-09-949-016-2936
; Sequence 2936, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2936
; LENGTH: 2001
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2936

Query Match          94.7%; Score 74.8; DB 3; Length 2001;
Best Local Similarity 97.4%; Pred. No. 4e-15;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCGCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
DB 1646 CGGGAATCCCGCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1705
QY 61 AACAAACCGACTCTAAAA 78
DB 1706 AACAAACCGACTCAAAA 1723

RESULT 6
US-08-755-559-2
```

; Sequence 2, Application US/08755559
; Patent No. 5912142
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,559
; FILING DATE: 22-NOV-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-116
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 2:
; LENGTH: 2180 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-755-559-2

Query Match 94.7%; Score 74.8; DB 2; Length 2180;
Best Local Similarity 97.4%; Pred. No. 4.1e-15;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1841 CGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1900

Qy 61 AACAAACCGACTCTAAAA 78
Db 1901 AACAAACCGACTCTACAA 1918

RESULT 7
US-09-210-474-2
; Sequence 2, Application US/09210474
; Patent No. 6072034
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,474
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,559
; FILING DATE: 22-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-116
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 2:
; LENGTH: 2180 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-210-474-2

Query Match 94.7%; Score 74.8; DB 3; Length 2180;
Best Local Similarity 97.4%; Pred. No. 4.1e-15;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1841 CGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1900

Qy 61 AACAAACCGACTCTAAAA 78
Db 1901 AACAAACCGACTCTACAA 1918

RESULT 8
US-09-539-774-2
; Sequence 2, Application US/09539774
; Patent No. 6350615
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/539,774
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/210,474
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-116
; TELEPHONE: (703) 816-4000

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14546
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-14546

Query Match 35.2%; Score 27.8; DB 3; Length 818128;
Best Local Similarity 65.1%; Pred. No. 50;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 12 CTTGAGTTCCTTGAAGAAAGTCCGACTCGAATATCTGAATGAAGAAACAAACCGAC 71
Db 61735 CTTGAGTTCCTGAGCACATCCGAGTTTCTTATCTCTGGAGTGCAGACTTCAAAAATAC 61794

Qy 72 TCT 74
Db 61795 TCT 61797

RESULT 13
US-09-949-016-14547
; Sequence 14547, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14547
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-14547

Query Match 35.2%; Score 27.8; DB 3; Length 818128;
Best Local Similarity 65.1%; Pred. No. 50;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 12 CTTGAGTTCCTTGAAGAAAGTCCGACTCGAATATCTGAATGAAGAAACAAACCGAC 71
Db 61735 CTTGAGTTCCTGAGCACATCCGAGTTTCTTATCTCTGGAGTGCAGACTTCAAAAATAC 61794

Qy 72 TCT 74
Db 61795 TCT 61797

RESULT 14
US-09-949-016-14548
; Sequence 14548, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14548
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-14548

Query Match 35.2%; Score 27.8; DB 3; Length 818128;
Best Local Similarity 65.1%; Pred. No. 50;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 12 CTTGAGTTCCTTGAAGAAAGTCCGACTCGAATATCTGAATGAAGAAACAAACCGAC 71
Db 61735 CTTGAGTTCCTGAGCACATCCGAGTTTCTTATCTCTGGAGTGCAGACTTCAAAAATAC 61794

Qy 72 TCT 74
Db 61795 TCT 61797

RESULT 15
US-09-949-016-14549
; Sequence 14549, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14549
; LENGTH: 818128
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; US-09-949-016-14549

```
; LOCATION: (1)...(818128)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14549

Query Match      35.2%; Score 27.8; DB 3; Length 818128;
Best Local Similarity 65.1%; Pred. No. 50;
Matches 41; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY      12 CTTTCAGTTCCTTGAAGAAGTTCATGACTCGAATATCTGAAATGAGAGAAACAAACCGAC 71
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      61735 CTTTCAGTTCCTGAGCAGACATTCGCGAGTTTCTTATCTCTCGAGTGCAGACTTCAAAATAC 61794
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY      72 TCT 74
        |||
Db      61795 TCT 61797
```

Search completed: December 6, 2005, 11:47:18
Job time : 44.4809 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 901.725 Seconds
(without alignments)
4099.008 Million cell updates/sec

Title: US-09-980-046B-3

Perfect score: 79

Sequence: 1 cgggaatccccccttcagttc.....aaacaaccgactctaaaaa 79

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gsl1:*
10: gb_gsl2:*
11: gb_gsl3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	75.4	96.7	479	5	EX108872
C 2	75.8	95.9	292	2	BF431778
C 3	75.8	95.9	313	3	BM053069
C 4	75.8	95.9	338	7	CK903980
C 5	75.8	95.9	339	7	CK903979
C 6	75.8	95.9	475	1	AI927624
C 7	75.8	95.9	633	3	BQ182988
C 8	75.8	95.9	684	6	CB241372
C 9	75.8	95.9	691	6	CD364677
C 10	75.8	95.9	716	2	BF002331
C 11	74.8	94.7	109	1	AA826896
C 12	74.8	94.7	180	1	AA874230
C 13	74.8	94.7	272	2	BF002304
C 14	74.8	94.7	310	1	AA903087
C 15	74.8	94.7	332	1	AI621135
C 16	74.8	94.7	376	1	AI281889
C 17	74.8	94.7	417	1	AW002382
C 18	74.8	94.7	418	1	AW594044
C 19	74.8	94.7	420	1	AI195936
C 20	74.8	94.7	429	1	AW207187
C 21	74.8	94.7	432	1	AW057702
C 22	74.8	94.7	440	1	AI218726

C 23	74.8	94.7	442	1	AI652267	AI652267 wb28d06.x
C 24	74.8	94.7	453	1	AI805092	AI805092 tu33g10.x
C 25	74.8	94.7	522	2	BF939675	BF939675 nac79g07.x
C 26	74.8	94.7	687	5	BU689727	BU689727 UI-CF-FN0
C 27	74.8	94.7	1633	4	CR595956	CR595956 full-leng
C 28	73.8	93.4	429	1	AW081554	AW081554 xc42g10.x
C 29	73.8	93.4	456	1	AI1979001	AI1979001 wr69a10.x
C 30	73	92.4	408	2	BE677472	BE677472 7d84g07.x
C 31	73	92.4	413	1	AI433486	AI433486 ti53b01.x
C 32	73	92.4	413	1	AI433494	AI433494 ti53c03.x
C 33	73	92.4	1650	4	CR594943	CR594943 full-leng
C 34	72.6	91.9	519	1	AW014948	AW014948 UI-H-B10-
C 35	72.6	91.9	693	6	CB048392	CB048392 NISC-9704
C 36	71.8	90.9	322	5	EX378127	EX378127 BX378127
C 37	71.6	90.6	350	1	AW974325	AW974325 EST386428
C 38	71.4	90.4	157	1	AI277084	AI277084 ql40h08.x
C 39	70.6	89.4	158	1	AI886304	AI886304 wm67a04.x
C 40	63.6	80.5	445	1	AI480347	AI480347 tm51c09.x
C 41	63.6	80.5	1064	5	EX335528	EX335528 BX335528
C 42	63.4	80.3	209	5	C01591	C01591 HUMG000859
C 43	62.2	78.7	799	5	EX380143	EX380143 BX380143
C 44	62	78.5	445	1	AW058447	AW058447 wx21a06.x
C 45	58	73.4	1700	4	CR614987	CR614987 full-leng

ALIGNMENTS

RESULT 1
EX108872/c

LOCUS

DEFINITION

IMAGE:1961756, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

EX108872 479 bp mRNA linear EST 07-FEB-2003
BX108872 NCI CGAP Col4 Homo sapiens CDNA clone IMAGp998A214818 ;
IMAGE:1961756, mRNA sequence.

ACCESSION
EX108872

VERSION
EX108872.1

KEYWORDS
GI:27835474

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 479)

AUTHORS
Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
Radelof,U., Schneider,D. and Korn,B.

TITLE
Human UnigeneSet - RZPD3

JOURNAL
Unpublished (2003)

COMMENT
Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Average insert size 1.7 kb. Life Technologies catalog #:
11531-019"

```

ORIGIN
Query Match          96.7%; Score 76.4; DB 5; Length 479;
Best Local Similarity 97.5%; Pred. No. 8e-13;
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCATGACTCGAATATCTCGAATGAAGAA 60
    |||||
Db 102 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCATGACTCGAATATCTCGAATGAAGAA 43
    |||||

Qy 61 AACAAACCGACTCTAAAAA 79
    |||||
Db 42 AACAAACCGACTCNACAAA 24
    |||||

RESULT 2
BF431778/c          292 bp      mRNA      linear      EST 19-JAN-2001
LOCUS
DEFINITION
nab50c04.x1 Soares NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
IMAGE:3269358 3', mRNA sequence.
ACCESSION
BF431778
VERSION
BF431778.1 GI:11443892
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 292)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cga@b6mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 272.
Location/Qualifiers
1..292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3269358"
/lab_host="DH10B"
/clone_lib="Soares NSF_F8_9W_OT_PA_P_S1"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NBHSF pool 1:
309384-310919, 323208-325895 Soares Nb2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares Nb2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NBHPA pool 1:
304776-326311, 320136-322823, 326280-326663 Soares NBHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN
Query Match          95.9%; Score 75.8; DB 2; Length 292;
Best Local Similarity 97.5%; Pred. No. 1.1e-12;
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCATGACTCGAATATCTCGAATGAAGAA 60
    |||||

```

```

Db 81 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCATGACTCGAATATCTCGAATGAAGAA 22

Qy 61 AACAAACCGACTCTAAAAA 79
    |||||
Db 21 AACAAACCGACTCACAAA 3
    |||||

RESULT 3
BM053069
LOCUS
DEFINITION
1e69h03.y3 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA clone IMAGE:5672308 5', mRNA sequence.
ACCESSION
BM053069
VERSION
BM053069.1 GI:16809019
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 313)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarisvill, R.,
Williams, T., Jackson, Y. and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center. For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu) This sequence now available from the IMAGE
consortium, for clone orders contact: info@image.llnl.gov.
Location/Qualifiers
1..313
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5672308"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/note="Organ: Pancreas; Vector: pSPORT1; Site 1: Not 1;
Site 2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dt priming. Size-selected by column
fractionation; average insert size 1.08 Kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an EcoT of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

FEATURES
source

```

```

Query Match          95.9%; Score 75.8; DB 3; Length 313;
Best Local Similarity 97.5%; Pred. No. 1.2e-12;
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```


Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
 |||
 Db 229 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 288
 |||
 Qy 61 AACAAACCGACTCTAATAA 79
 |||
 Db 289 AACAAACCGACTCACAATA 307
 |||

RESULT 4
 CK903980
 LOCUS ie69h03.y5 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
 DEFINITION cDNA clone IMAGE:5672308 5', mRNA sequence.
 CK903980
 ACCESSION CK903980.1 GI:45365511
 VERSION
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 338)
 Melton,D., Meadows,A., Clifton,S., Hillier,L., Marra,M., Pape,D.,
 Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B.,
 Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M.,
 McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and
 Bowers,Y.

REFERENCE
 AUTHORS WashU-Harvard Pancreas EST Project
 TITLE Unpublished (2000)
 JOURNAL Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 COMMENT Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138

Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohp.harvard.edu
 This read is a 5' RESEQUENCE of a previously sequenced pancreas
 clone
 This read has been verified (found to hit its original self in the
 correct orientation)
 Putative full length read
 vector to vector length is
 vector to vector length is
 Seq primer: -40RP from Gibco.
 Location/Qualifiers
 1..338
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5672308"
 /sex="Both"
 /tissue_type="Islets of Langerhans"
 /dev_stage="Adult"
 /lab_host="DH10B"
 /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
 /note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
 Site_2: Sal 1; Starting library constructed using
 SuperScript Plasmid Library kit (Life Technologies). cDNA
 made by oligo-dT priming. Size-selected by column
 fractionation; average insert size 1.08 kb. Library was
 amplified once on solid support and plasmid DNA from
 library was prepared. The library DNA was normalized by
 method #4 from Bonaldo, Lennon, and Soares 1996 Genome
 Research 6:791-806; 0.5 microgram single-stranded library
 plasmid DNA was mixed with 5 micrograms PCR product
 representing library inserts and hybridized to an Ecot of
 20. Single-stranded (unhybridized) plasmids were isolated
 by hydroxyapatite chromatography and used to make this
 library."

FEATURES

source

FEATURES

source

1..339
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5672308"
 /sex="Both"
 /tissue_type="Islets of Langerhans"
 /dev_stage="Adult"
 /lab_host="DH10B"
 /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
 /note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
 Site_2: Sal 1; Starting library constructed using
 SuperScript Plasmid Library kit (Life Technologies). cDNA
 made by oligo-dT priming. Size-selected by column
 fractionation; average insert size 1.08 kb. Library was
 amplified once on solid support and plasmid DNA from
 library was prepared. The library DNA was normalized by
 method #4 from Bonaldo, Lennon, and Soares 1996 Genome
 Research 6:791-806; 0.5 microgram single-stranded library
 plasmid DNA was mixed with 5 micrograms PCR product
 representing library inserts and hybridized to an Ecot of
 20. Single-stranded (unhybridized) plasmids were isolated
 by hydroxyapatite chromatography and used to make this
 library."

ORIGIN

Query Match 95.9%; Score 75.8; DB 7; Length 338;

ORIGIN

Best Local Similarity 97.5%; Pred. No. 1.2e-12;
 Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
 |||
 Db 247 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 306
 |||
 Qy 61 AACAAACCGACTCTAATAA 79
 |||
 Db 307 AACAAACCGACTCACAATA 325
 |||

RESULT 5

CK903979/c

LOCUS

ie69h03.x5 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens

DEFINITION cDNA clone IMAGE:5672308 3', mRNA sequence.

CK903979

ACCESSION CK903979.1 GI:45365510

VERSION

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 339)
 Melton,D., Meadows,A., Clifton,S., Hillier,L., Marra,M., Pape,D.,
 Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B.,
 Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M.,
 McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and
 Bowers,Y.

REFERENCE

AUTHORS WashU-Harvard Pancreas EST Project

TITLE Unpublished (2000)

JOURNAL Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

COMMENT Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

This read is a 3' RESEQUENCE of a previously sequenced pancreas
 clone
 This resequenced clone has not previously been sequenced on this
 end, resequencing from this end represents new data
 Seq primer: -400P from Gibco.
 Location/Qualifiers
 1..339
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5672308"
 /sex="Both"
 /tissue_type="Islets of Langerhans"
 /dev_stage="Adult"
 /lab_host="DH10B"
 /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
 /note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
 Site_2: Sal 1; Starting library constructed using
 SuperScript Plasmid Library kit (Life Technologies). cDNA
 made by oligo-dT priming. Size-selected by column
 fractionation; average insert size 1.08 kb. Library was
 amplified once on solid support and plasmid DNA from
 library was prepared. The library DNA was normalized by
 method #4 from Bonaldo, Lennon, and Soares 1996 Genome
 Research 6:791-806; 0.5 microgram single-stranded library
 plasmid DNA was mixed with 5 micrograms PCR product
 representing library inserts and hybridized to an Ecot of
 20. Single-stranded (unhybridized) plasmids were isolated
 by hydroxyapatite chromatography and used to make this
 library."

Query Match 95.9%; Score 75.8; DB 7; Length 339;
 Best Local Similarity 97.5%; Pred. No. 1.2e-12;
 Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAATCCCTTCAGTCTTTTGAAGAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 Db 81 CGGAATCCCTTCAGTCTTTTGAAGAAGTTCATGACTCGAATATCTGAAATGAAGAA 22
 |||||

Qy 61 AACAAACCGACTCTAAAAA 79
 |||||
 Db 21 AACAAACCGACTCTAAAAA 3

RESULT 6
 A1927624/c
 LOCUS w085a02.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2462090 3',
 DEFINITION mRNA sequence.

ACCESSION A1927624
 VERSION A1927624.1 GI:5663588

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE 1 (bases 1 to 475)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Straubeberg, Ph.D.

Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
 Clone Distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 1810 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 456.

FEATURES

source

1..475

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2462090"

/lab_host="DH10B"

/clone_lib="NCI_CGAP Kid11"

/note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with

a modified polylinker; Site 1: Not I; Site 2: Eco RI;

Plasmid DNA from the normalized library NCI_CGAP Kid3 was

prepared, and ss circles were made in vitro. Following RAP

purification, this DNA was used as tracer in a subtractive

hybridization reaction. The driver was PCR-amplified cDNAs

from a pool of 5,000 clones made from the same library

(cloneIDs 1322376-1323911, 1456007-1456775, and

1500552-1502855). Subtraction by Bento Soares and M.

Fatima Bonaldo."

ORIGIN

Query Match 95.9%; Score 75.8; DB 1; Length 475;

Best Local Similarity 97.5%; Pred. No. 1.2e-12;

Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAATCCCTTCAGTCTTTTGAAGAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 Db 81 CGGAATCCCTTCAGTCTTTTGAAGAAGTTCATGACTCGAATATCTGAAATGAAGAA 22
 |||||

Qy 61 AACAAACCGACTCTAAAAA 79
 |||||
 Db 21 AACAAACCGACTCTAAAAA 3

RESULT 7
 BQ182988/c
 LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE 1 (bases 1 to 633)

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Straubeberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Jose Mercuende

CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

The following repetitive elements were found in this cDNA

sequence: 276-341 >GC richLow_complexity

Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES

source

1..633

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5839650"

/tissue_type="Chondrosarcoma"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI_CGAP ED1"

/note="Organ: Left Pubic Bone; Vector: pT7T3-Pac

(Pharmacia) with a modified polylinker; Site 1: EcoR I;

Site 2: Not I; NCI_CGAP ED1 is a normalized cDNA library

containing the following tissue(s): Chondrosarcoma cell

line C55. The library was constructed according to

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. First strand cDNA synthesis was primed with an

oligo-dT primer containing a Not I site. Double stranded

cDNA was ligated to an EcoR I adaptor, digested with Not

I, and cloned directionally into pT7T3-Pac vector. The

oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is

located between the Not I site and the (dT)18 tail. The

sequence tag for this library is GCTCAAGGCT.

TAG TISSUE=chondrosarcoma

TAG LIB=UI-H-ED1

TAG_SEQ=CGTCAAGGCT"

Query Match 95.9%; Score 75.8; DB 3; Length 633;

Best Local Similarity 97.5%; Pred. No. 1.3e-12;

Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAATCCCTTCAGTCTTTTGAAGAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 Db 91 CGGAATCCCTTCAGTCTTTTGAAGAAGTTCATGACTCGAATATCTGAAATGAAGAA 32
 |||||

Qy 61 AACAAACCGACTCTAAAAA 79

Db	31	32
Db	31 AACAAACCGACTCACA AAA 13	32 AACAAACCGACTCACA AAA 14
RESULT 8	CB241372	CD364677
LOCUS	UI-CF-FNO-afy-h-23-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone	UI-H-FT2-bjm-1-14-0-UI.s1 NCI CGAP FT2 Homo sapiens cDNA clone
DEFINITION	UI-CF-FNO-afy-h-23-0-UI 3', mRNA sequence.	UI-H-FT2-bjm-1-14-0-UI 3', mRNA sequence.
ACCESSION	CB241372	CD364677
VERSION	CB241372.1 GI:28363016	CD364677.1 GI:31148767
KEYWORDS	EST.	EST.
SOURCE	Homo sapiens (human)	Homo sapiens (human)
ORGANISM	Homo sapiens	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 684)	1 (bases 1 to 691)
TITLE	Bonaldo, M.F., Lennon, G. and Soares, M.B.	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL	Normalization and subtraction: two approaches to facilitate gene discovery	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
PUBMED	Genome Res. 6 (9), 791-806 (1996)	Unpublished (1997)
COMMENT	889548 Contact: McCray, PB McCray Lab University of Iowa 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA Tel: 319 356 4866 Fax: 319 356 7171 Email: paul-mccray@uiowa.edu Tissue Procurement: Dr. M. J. Welsh, University of Iowa cDNA Library preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Researchers may obtain clones from Research Genetics (www.regen.com) or from Open Biosystems (www.openbiosystems.com). The following repetitive elements were found in this cDNA sequence: 277-342, >GC rich#Low_complexity Seq primer: M13 FORWARD POLYA=Yes.	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Dr. Gary W. Hunninghake, U of I cDNA Library preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/cgap.html Seq primer: M13 FORWARD POLYA=Yes.
FEATURES	Location/Qualifiers 1..684 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="UI-CF-FNO-afy-h-23-0-UI" /tissue_type="Human Lung Epithelial cells" /lab_host="DH10B (Life Technologies)" /clone_lib="UI-CF-FNO" /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-FNO is a subtracted cDNA library derived from two normalized Human lung epithelial cell libraries (EN1 and DU1) The library was subtracted according to according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bento-soares@uiowa.edu TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS 6hr to LPS 24hr TAG LIB=UI-CF-FNO TAG_SEQ=CTGCTCAGGT"	Location/Qualifiers 1..691 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="UI-H-FT2-bjm-1-14-0-UI" /tissue_type="Alveolar Macrophage" /dev_stage="Adult" /lab_host="DH10B (Life Technologies)" /clone_lib="NCI CGAP FT2" /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FT2 is a subtracted cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The mRNA samples were a mixture of these conditions (times refer to incubations following isolation by bronchoalveolar lavage) (some normal donor macrophages were cultured in some of the conditions, other donor macrophages in different conditions). The mRNA samples were pooled for library construction. Control 0 hours; control 3 hours; control 24 hours; LPS 100 ng/ml, 3 hours; LPS 100 ng/ml, 24 hours; PMA 10 ng/ml, 3 hours; PMA 10 ng/ml, 24 hours; Klebsiella moi 10, 3 hours; Klebsiella moi 10, 24 hours; Staph aureus moi 10, 3 hours; Staph aureus moi 10, 24 hours; Adenoviral vector (Ad5 CMV eGFP), moi 500, 3 hours; Adenoviral vector (Ad5 CMV eGFP), moi 500, 24 hours; wt adenovirus moi 500, 3 hours; wt adenovirus moi 500, 24 hours; Ad vector + LPS 3 hours; Ad vector + LPS 24 hours; wt adenovirus + LPS 3 hours; wt adenovirus + LPS 24 hours. The library was subtracted according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa. TAG_TISSUE=Human Lung Alveolar Macrophage TAG LIB=UI-H-FT2 TAG_SEQ=GGCCATGCGC"
ORIGIN	Query Match 95.9%; Score 75.8; DB 6; Length 684; Best Local Similarity 97.5%; Pred. No. 1.3e-12; Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	Query Match 95.9%; Score 75.8; DB 6; Length 691; Best Local Similarity 97.5%; Pred. No. 1.3e-12; Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy	1 CGGGAATCCCTTCAGTTCCTTTTCAAAAAGTTCCATGCTCGAATATCTGAAATGAAGAA 60	1 CGGGAATCCCTTCAGTTCCTTTTCAAAAAGTTCCATGCTCGAATATCTGAAATGAAGAA 60
Db	92 CGGGAATCCCTTCAGTTCCTTTTCAAAAAGTTCCATGCTCGAATATCTGAAATGAAGAA 33	92 CGGGAATCCCTTCAGTTCCTTTTCAAAAAGTTCCATGCTCGAATATCTGAAATGAAGAA 33
Oy	61 AACAAACCGACTCATAAAA 79	61 AACAAACCGACTCATAAAA 79

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 180)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Klaus Kaserer, M.D., Chris Moskaluk, M.D.,
Ph.D., Michael R. Emmert-Buck, M.D., P.H.D.
cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov

Seq primer: -40UP from Gibco

High quality sequence stop: 168.

FEATURES
source

1..180
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3126297"
/tissue_type="follicular carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NCI CGAP Thy3"
/note="Organ: thyroid; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library constructed by Life
Technologies."

Query Match 94.7%; Score 74.8; DB 1; Length 180;

Best Local Similarity 97.4%; Pred. No. 2.2e-12;

Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 60

Db 85 CGGGAATCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 26

Qy 61 AACAAACCGACTCTAAAA 78

Db 25 AACAAACCGACTCTAAAA 8

RESULT 13

LOCUS BF002304/C

DEFINITION 7h02904.x1 NCI CGAP Col6 Homo sapiens cDNA clone IMAGE:3314838 3',

Accession BF002304

Version BF002304.1 GI:10702579

Keywords EST.

Source Homo sapiens (human)

Organism Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 272)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck,

M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov

Seq primer: -40UP from Gibco.

FEATURES
source

1..272
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3314838"
/tissue_type="colon tumor, RER+"
/lab_host="DH10B"
/clone_lib="NCI CGAP Col6"

/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Col0 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1057416-1061255, and 1144584-1145351).
Subtraction by Bento Soares and M. Fatima Bonaldo."

Query Match 94.7%; Score 74.8; DB 2; Length 272;

Best Local Similarity 97.4%; Pred. No. 2.3e-12;

Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 60

Db 83 CGGGAATCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 24

Qy 61 AACAAACCGACTCTAAAA 78

Db 28 AACAAACCGACTCTAAAA 6

RESULT 14

LOCUS AA903087/C

DEFINITION Q46d04.s1 NCI CGAP Lei2 Homo sapiens cDNA clone IMAGE:1516999 3',

Accession AA903087

Version AA903087.1 GI:3038210

Keywords EST.

Source Homo sapiens (human)

Organism Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 310)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Unknown library type

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 295.

Location/Qualifiers

1..310
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1516999"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B"
/clone_lib="NCI CGAP Lei2"

/note="Organ: soft tissue; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Site 1: Not I; Site 2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
[5' -AACTGGAAGAATTCGGCGCGCAATCGTTTTTTTTTTTTTTTTTT-3'],

double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTV3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 94.7%; Score 74.8; DB 1; Length 310;
 Best Local Similarity 97.4%; Pred. No. 2.3e-12;
 Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 Db 90 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 31
 |||||

QY 61 AACAAACCGACTCTAAAA 78
 |||||
 Db 30 AACAAACCGACTCAGAAA 13
 |||||

RESULT 15

AI621135/c
 LOCUS
 DEFINITION tu49d01.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2254369 3',
 mRNA sequence.
 ACCESSION AI621135
 VERSION AI621135.1 GI:4630261
 KEYWORDS EST.
 SOURCE
 ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE

1 (bases 1 to 332)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 582 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 320

POLYA=No.

FEATURES

source

1. .332
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2254369"
 /sex="male"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_Pr28"

/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia)
 with a modified polylinker; Plasmid DNA from the
 normalized library NCI CGAP Pr22 was prepared, and ss
 circles were made in vitro. Following HAP purification,
 this DNA was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from a pool
 of 5,000 clones made from the same library (clones
 985608-986759, 1101192-1101959, and 1217928-1220615).
 Subtraction by Bento Soares and M. Fatima Bonaldo. "

ORIGIN

Query Match 94.7%; Score 74.8; DB 1; Length 332;
 Best Local Similarity 97.4%; Pred. No. 2.4e-12;
 Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 Db 131 CGGGAATCCCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 72
 |||||

QY 61 AACAAACCGACTCTAAAA 78
 |||||
 Db 71 AACAAACCGACTCAGAAA 54
 |||||

Search completed: December 7, 2005, 09:55:48

Job time : 904.825 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 38.8456 Seconds
(without alignments)
4072.612 Million cell updates/sec

Title: US-09-980-046B-4
Perfect score: 89
Sequence: 1 cgggaatcccccttcagtc.....aaaaaaaaaaaaaagaaaa 89

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 130357 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*

5: /cgn2_6/ptodata/1/ina/H COMB.seq:*

6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*

7: /cgn2_6/ptodata/1/ina/PP COMB.seq:*

8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*

9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	70	78.7	601	3	US-09-949-016-26334, A
C 2	70	78.7	601	3	US-09-949-016-105203, A
C 3	70	78.7	2000	3	US-09-997-165-3, Appli
4	70	78.7	2000	3	US-09-949-016-426, App
5	70	78.7	2001	3	US-09-949-016-2936, App
6	70	78.7	2180	2	US-08-755-559-2, Appli
7	70	78.7	2180	2	US-09-210-474-2, Appli
8	70	78.7	2180	3	US-09-539-774-2, Appli
9	70	78.7	16738	3	US-09-949-016-12168, A
10	70	78.7	16738	3	US-09-949-016-14678, A
11	43.6	49.0	177	3	US-09-621-976-8073, Ap
12	41	46.1	164	3	US-09-621-976-8070, Ap
13	41	46.1	166	3	US-09-621-976-8651, Ap
14	40.6	45.6	1477	3	US-09-585-173B-7, Appli
15	40.4	45.4	246	3	US-09-621-976-16288, A
16	40.4	45.4	259	3	US-09-621-976-16294, A
17	39.4	44.3	249	3	US-09-621-976-16291, A
18	38.6	43.4	249	3	US-09-621-976-16292, A
19	38.6	43.4	732	3	US-09-149-476-66, Appli
20	38.6	43.4	1117	3	US-09-247-373B-33, Appli
21	38.6	43.4	2091	3	US-09-813-818-1, Appli
22	38.6	43.4	2091	3	US-10-199-333-1, Appli
23	38.2	42.9	127	3	US-09-621-976-13933, A
24	38	42.7	159	3	US-09-621-976-8623, Ap

25	38	42.7	159	3	US-09-621-976-8655	Sequence 8655, Ap
26	37.8	42.5	1743	3	US-10-012-231A-235	Sequence 235, App
27	37.8	42.5	1743	3	US-10-015-389A-235	Sequence 235, App
28	37.8	42.5	1743	3	US-10-006-768A-235	Sequence 235, App
29	37.8	42.5	1743	3	US-10-015-671A-235	Sequence 235, App
30	37.8	42.5	1743	3	US-10-015-393A-235	Sequence 235, App
31	37.8	42.5	1743	3	US-10-011-833A-235	Sequence 235, App
32	37.8	42.5	1743	3	US-10-006-041A-235	Sequence 235, App
33	37.8	42.5	1743	3	US-10-012-064A-235	Sequence 235, App
34	37.4	42.0	266	3	US-09-621-976-16813	Sequence 16813, A
35	37.4	42.0	3275	3	US-09-370-838-151	Sequence 151, App
36	37.4	42.0	3275	3	US-09-854-133-151	Sequence 151, App
C 37	37.2	41.8	11050	3	US-09-949-016-14116	Sequence 14116, A
C 38	37	41.6	194	3	US-10-131-827-8633	Sequence 8633, Ap
39	36.8	41.3	1882	3	US-09-419-679-3	Sequence 3, Appli
40	36.6	41.1	1406	3	US-10-000-489-81	Sequence 81, Appli
41	36.6	41.1	1934	3	US-08-776-844-1	Sequence 1, Appli
42	36.6	41.1	1934	3	US-09-909-325-1	Sequence 1, Appli
43	36.6	41.1	1934	3	US-09-909-326-1	Sequence 1, Appli
44	36.6	41.1	5520	3	US-10-001-887-43	Sequence 43, Appli
C 45	36.4	40.9	226	3	US-10-131-827-8671	Sequence 8671, Ap

ALIGNMENTS

RESULT 1
US-09-949-016-26334/c
; Sequence 26334, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26334
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-26334

Query Match 78.7%; Score 70; DB 3; Length 601;
Best Local Similarity 93.6%; Pred. No. 1.6e-08;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy	1	CGGGAATCCCCTTCAGTCTTTTGAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db	129	CGGGAATCCCCTTCAGTCTTTTGAAGTTCCATGACTCGAATATCTGAAATGAAGAA 70
Oy	61	AACAAACCAAAAAA 78
Db	69	AACAAACCGACTCANA 52

RESULT 2
US-09-949-016-105203/c
; Sequence 105203, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105203
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-105203

Query Match          78.7%; Score 70; DB 3; Length 601;
Best Local Similarity 93.6%; Pred. No. 1.6e-08;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 129 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 70

Qy 61 AACAAACCAAAAAA 78
Db 69 AACAAACCGACTCACA 52

RESULT 3
US-09-997-165-3
; Sequence 3, Application US/09997165
; Patent No. 6762030
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; TITLE OF INVENTION: FANBLOW, WILLIAM C.
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 3
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)..(865)
US-09-997-165-3

Query Match          78.7%; Score 70; DB 3; Length 2000;
Best Local Similarity 93.6%; Pred. No. 1.7e-08;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Db 1705 AACAAACCGACTCACA 1722

RESULT 4
US-09-949-016-426
; Sequence 426, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 426
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-426

Query Match          78.7%; Score 70; DB 3; Length 2000;
Best Local Similarity 93.6%; Pred. No. 1.7e-08;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Db 1705 AACAAACCGACTCACA 1722

RESULT 5
US-09-949-016-2936
; Sequence 2936, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2936
; LENGTH: 2001
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2936

Query Match          78.7%; Score 70; DB 3; Length 2001;
Best Local Similarity 93.6%; Pred. No. 1.7e-08;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1646 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1705

Qy 61 AACAAACCAAAAAA 78
Db 1706 AACAAACCGACTCACA 1723

RESULT 6
US-08-755-559-2
```


; Sequence 2, Application US/0875559
; Patent No. 5912142
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,559
; FILING DATE: 22-NOV-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2180 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-755-559-2

Query Match 78.7%; Score 70; DB 2; Length 2180;
Best Local Similarity 93.6%; Pred. No. 1.7e-08;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db |||||
Qy 61 AACAAACCAAAAAA 78
Db |||||

RESULT 7
US-09-210-474-2
; Sequence 2, Application US/09210474
; Patent No. 6072034
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/539,774
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/210,474
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,474
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,559
; FILING DATE: 22-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2180 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-210-474-2

Query Match 78.7%; Score 70; DB 3; Length 2180;
Best Local Similarity 93.6%; Pred. No. 1.7e-08;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db |||||
Qy 61 AACAAACCAAAAAA 78
Db |||||

RESULT 8
US-09-539-774-2
; Sequence 2, Application US/09539774
; Patent No. 6350615
; GENERAL INFORMATION:
; APPLICANT: KAUFMAN, RUSSEL E.
; APPLICANT: SLENTZ-KESLER, KIMBERLY
; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
; TITLE OF INVENTION: CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/539,774
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/210,474
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000

```
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 2180 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
US-09-539-774-2

Query Match      78.7%; Score 70; DB 3; Length 2180;
Best Local Similarity 93.6%; Pred. No. 1.7e-08;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1841 CGGGAATCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1900

Qy 61 AACAAACCAAAAAAAAAA 78
Db 1901 AACAAACCGACTCAGAA 1918

RESULT 9
US-09-949-016-12168
; Sequence 12168, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12168
; LENGTH: 16738
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12168

Query Match      78.7%; Score 70; DB 3; Length 16738;
Best Local Similarity 93.6%; Pred. No. 2e-08;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 14383 CGGGAATCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 14442

Qy 61 AACAAACCAAAAAAAAAA 78
Db 14443 AACAAACCGACTCAGAA 14460

RESULT 10
US-09-949-016-14678
; Sequence 14678, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; SOFTWARE: Patent.pm

Query Match      78.7%; Score 70; DB 3; Length 16738;
Best Local Similarity 93.6%; Pred. No. 2e-08;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 14383 CGGGAATCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 14442

Qy 61 AACAAACCAAAAAAAAAA 78
Db 14443 AACAAACCGACTCAGAA 14460

RESULT 11
US-09-621-976-8073
; Sequence 8073, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8073
; LENGTH: 177
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8073

Query Match      49.0%; Score 43.6; DB 3; Length 177;
Best Local Similarity 69.0%; Pred. No. 0.04;
Matches 58; Conservative 1; Mismatches 25; Indels 0; Gaps 0;

Qy 6 ATCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAAACAA 65
Db 80 ATCTCTCTCCCATTTTGAAGAAAGTTTCAATAGAAAAAAGAAAAAAGAAAAAAGAAAA 139

Qy 66 ACCAAAAAAGAAAAAAGAAAAA 89
Db 140 AAAAAAAGAAAAAAGAAAAAAGAAAAA 163

RESULT 12
US-09-621-976-8070
; Sequence 8070, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
```

```
; SEQ ID NO 8070
; LENGTH: 164
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8070

Query Match          46.1%; Score 41; DB 3; Length 164;
Best Local Similarity 69.1%; Pred. No. 0.17;
Matches 56; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 9 CCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAAAAACAACC 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 82 CCTCCTCCATTTTGTGAATAGTTTCAATAGAAAAAAGAAAAAAGAAAAAAGAAAAA 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 69 AAAAAAAGAAAAAAGAAAAA 89
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 142 AAAAAAAGAAAAAAGAAAAA 162
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-09-621-976-8651
; Sequence 8651, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8651
; LENGTH: 166
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8651

Query Match          46.1%; Score 41; DB 3; Length 166;
Best Local Similarity 69.1%; Pred. No. 0.17;
Matches 56; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 9 CCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAAAAACAACC 68
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 82 CCTCCTCCATTTTGTGAATAGTTTCAATAGAAAAAAGAAAAAAGAAAAAAGAAAAA 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 69 AAAAAAAGAAAAAAGAAAAA 89
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 142 AAAAAAAGAAAAAAGAAAAA 162
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-09-585-173B-7
; Sequence 7, Application US/09585173B
; Patent No. 6570063
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Gutteridge, Steven
; APPLICANT: Maxwell, Carl
; TITLE OF INVENTION: Magnesium Chelataase
; FILE REFERENCE: BB1370 US NA
; CURRENT APPLICATION NUMBER: US/09/585,173B
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/137,461
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 1477
; TYPE: DNA
; ORGANISM: Oryza sativa
```

```
US-09-585-173B-7
Query Match          45.6%; Score 40.6; DB 3; Length 1477;
Best Local Similarity 69.6%; Pred. No. 0.25;
Matches 55; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 11 CCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAATGAAGAAAAACAACC 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1379 CGTACAGTCTTAAGTAATAATAACATTCATAATTTATGTGAAAAAAGAAAAAAG 1438
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 71 AAAAAAAGAAAAAAGAAAAA 89
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1439 AAAAAAAGAAAAAAGAAAAA 1457
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
US-09-621-976-16288
; Sequence 16288, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16288
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16288

Query Match          45.4%; Score 40.4; DB 3; Length 246;
Best Local Similarity 75.8%; Pred. No. 0.25;
Matches 50; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 24 GAAAAAGTTCATGACTCGAATATCTGAATGAAGAAAAACAACCACAAAAAAGAAAA 83
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164 GAAAAAGCGGTAACTAATAGCTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 223
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 84 GAAAAA 89
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 AAAAAA 229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Search completed: December 6, 2005, 11:47:18
Job time : 38.8456 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 128.321 Seconds
(without alignments)
4072.612 Million cell updates/sec

Title: US-09-980-046B-7
Perfect score: 294
Sequence: 1 cggcgatcgcgcggttat.....agcgagcacctagacaaaaa 294

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	269.4	91.6	2369	2	US-08-883-534-4
2	269.4	91.6	2369	3	US-09-204-764-4
3	269.4	91.6	2921	3	US-09-949-016-1211
4	269.4	91.6	2921	3	US-09-949-016-1212
5	269.4	91.6	46492	3	US-09-949-016-12953
6	269.4	91.6	46492	3	US-09-949-016-12954
7	269	91.5	601	3	US-09-949-016-42279
8	269	91.5	601	3	US-09-949-016-42346
9	35.8	12.2	601	3	US-09-949-016-108017
10	35.8	12.2	32104	3	US-09-949-016-14722
11	31.4	10.7	12128	3	US-09-949-016-12587
12	31.4	10.7	12129	3	US-09-949-016-15713
13	30.8	10.5	1809	3	US-09-489-039A-5265
14	30.6	10.4	150	2	US-08-513-846-14
15	30.6	10.4	380	3	US-09-385-982-453
16	30.6	10.4	407	3	US-09-385-982-519
17	30.2	10.3	76563	3	US-09-949-016-17099
18	30.2	10.3	78269	3	US-09-949-016-12497
19	30	10.2	263	3	US-09-513-998C-34198
20	30	10.2	229354	3	US-09-705-400-64
21	30	10.2	462589	3	US-09-949-016-12900
22	30	10.2	476044	3	US-09-949-016-12412
23	29.6	10.1	2246	3	US-10-104-047-822
24	29.6	10.1	22118	3	US-09-815-981A-5

c 25 29.6 10.1 28257 3 US-09-949-016-13076
c 26 29.6 10.1 28438 3 US-09-820-790B-3
c 27 29.6 10.1 33112 3 US-10-429-873A-3
c 28 29.6 10.1 42693 3 US-09-949-016-17317
c 29 29.6 10.1 42693 3 US-09-949-016-17318
c 30 29.6 10.0 27727 3 US-09-949-016-15737
c 31 29.2 9.9 601 3 US-09-949-016-118858
c 32 29.2 9.9 601 3 US-09-949-016-118859
c 33 29.2 9.9 601 3 US-09-949-016-121206
c 34 29.2 9.9 601 3 US-09-949-016-121207
c 35 29.2 9.9 2007 3 US-09-620-312D-238
c 36 29.2 9.9 23218 3 US-09-949-016-11987
c 37 29.2 9.9 23219 3 US-09-949-016-13396
c 38 29.2 9.9 61158 3 US-09-949-016-15041
c 39 29.2 9.9 70563 3 US-09-949-016-16743
c 40 29.2 9.9 75431 3 US-09-949-016-15122
c 41 29 9.9 5538 3 US-09-949-016-14398
c 42 28.8 9.8 434 3 US-09-513-999C-10571
c 43 28.8 9.8 1641 3 US-09-375-932A-2
c 44 28.8 9.8 1641 3 US-10-156-562A-2
c 45 28.8 9.8 169998 3 US-09-676-610B-24

ALIGNMENTS

RESULT 1
US-08-883-534-4
; Sequence 4, Application US/08883534
; Patent No. 5846777
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: TWO NEW WD-40 PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/883,534
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0332 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2369 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: NEUTG01
; CLONE: 1221143
US-08-883-534-4

Sequence 13076, A
Sequence 3, Appli
Sequence 3, Appli
Sequence 17317, A
Sequence 17318, A
Sequence 15737, A
Sequence 118858,
Sequence 118859,
Sequence 121206,
Sequence 121207,
Sequence 238, App
Sequence 13987, A
Sequence 13396, A
Sequence 15041, A
Sequence 16743, A
Sequence 15122, A
Sequence 14398, A
Sequence 10571, A
Sequence 2, Appli
Sequence 2, Appli
Sequence 24, Appli

Query Match 91.6%; Score 269.4; DB 2; Length 2369;
Best Local Similarity 99.6%; Pred. No. 2.8e-80;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 21 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 80
Db 1898 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 1957

Qy 81 CTGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 140
Db 1958 CTGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 2017

Qy 141 TCTGTGACGGCGCCCATGCCCCCACCACCAAGAGGAGGAGGCGCCAGTCATGAC 200
Db 2018 TCTGTGACGGCGCCCATGCCCCCACCACCAAGAGGAGGAGGCGCCAGTCATGAC 2077

Qy 201 CCTCGTCTCTGACGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db 2078 CCTCGTCTCTGACGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 2137

Qy 261 CACATGAAAAATAAAGCGAGCACCTAGACAA 291
Db 2138 CACATGAAAAATAAAGCGAGCACCTAAACAA 2168

RESULT 2
US-09-204-764-4
; Sequence 4, Application US/09204764
; Patent No. 6025464
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: TWO NEW WD-40 PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/204,764
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/883,534
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0332 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2369 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: NEUTGMT01
; CLONE: 1221143
; US-09-204-764-4

Query Match 91.6%; Score 269.4; DB 3; Length 2369;
Best Local Similarity 99.6%; Pred. No. 2.8e-80;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 21 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 80
Db 1898 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 1957

Qy 81 CTGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 140
Db 1958 CTGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 2017

Qy 141 TCTGTGACGGCGCCCATGCCCCCACCACCAAGAGGAGGAGGCGCCAGTCATGAC 200
Db 2018 TCTGTGACGGCGCCCATGCCCCCACCACCAAGAGGAGGAGGCGCCAGTCATGAC 2077

Qy 201 CCTCGTCTCTGACGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db 2078 CCTCGTCTCTGACGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 2137

Qy 261 CACATGAAAAATAAAGCGAGCACCTAGACAA 291
Db 2138 CACATGAAAAATAAAGCGAGCACCTAAACAA 2168

RESULT 3
US-09-949-016-1211
; Sequence 1211, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1211
; LENGTH: 2921
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-1211

Query Match 91.6%; Score 269.4; DB 3; Length 2921;
Best Local Similarity 99.6%; Pred. No. 3e-80;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 21 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 80
Db 1982 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 2041

Qy 81 CTGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 140
Db 2042 CTGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 2101

Qy 141 TCTGTGACGGCGCCCATGCCCCCACCACCAAGAGGAGGAGGCGCCAGTCATGAC 200
Db 2102 TCTGTGACGGCGCCCATGCCCCCACCACCAAGAGGAGGAGGCGCCAGTCATGAC 2161

Qy 201 CCTCGTCTCTGACGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db 2162 CCTCGTCTCTGACGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 2221

Qy 261 CACATGAAAAATAAAGCGAGCACCTAGACAA 291
Db 2138 CACATGAAAAATAAAGCGAGCACCTAAACAA 2168

Db 2222 CACATGAAAAATAAAGCGAGCACCTAAACAA 2252

RESULT 4

US-09-949-016-1212
; Sequence 1212, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1212
; LENGTH: 2921
; TYPE: DNA
; ORGANISM: Human

US-09-949-016-1212

Query Match 91.6%; Score 269.4; DB 3; Length 2921;
Best Local Similarity 99.6%; Pred. No. 3e-80;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 21 TCCCATGATGCTCTGTCAAGGAGTGGAACAATCACCTACTGAGGAGCCACCCCGCCT 80
Db 1982 TCCCATGATGCTCTGTCAAGGAGTGGAACAATCACCTACTGAGGAGCCACCCCGCCT 2041

QY 81 CTGGATGACCGAATCAGGAGTGTCTGTACAGCTTTTAACTGCAGCGGAACATGTCATTTCTCTATT 140
Db 2042 CTGGATGACCGAATCAGGAGTGTCTGTACAGCTTTTAACTGCAGCGGAACATGTCATTTCTCTATT 2101
QY 141 TCTGTGACGGCCCCCATGCCCCCACCACCAAGAGGCGAGGAGCCCGCATGATGAC 200
Db 2102 TCTGTGACGGCCCCCATGCCCCCACCACCAAGAGGCGAGGAGCCCGCATGATGAC 2161
QY 201 CCTCGTCTCTGCAGGGTGTCTGTACAGCTTTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db 2162 CCTCGTCTCTGCAGGGTGTCTGTACAGCTTTTCTGAAAGCTTTAGACAGTAACAGTTTG 2221
QY 261 CACATGAAAAATAAAGCGAGCACCTAGACAA 291
Db 2222 CACATGAAAAATAAAGCGAGCACCTAAACAA 2252

RESULT 5

US-09-949-016-12953
; Sequence 12953, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12953

; LENGTH: 46492

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(46492)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-12953

Query Match 91.6%; Score 269.4; DB 3; Length 46492;
Best Local Similarity 99.6%; Pred. No. 9.1e-80;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 TCCCATGATGCTCTGTCAAGGAGTGGAACAATCACCTACTGAGGAGCCACCCCGCCT 80
Db 43553 TCCCATGATGCTCTGTCAAGGAGTGGAACAATCACCTACTGAGGAGCCACCCCGCCT 43612
QY 81 CTGGATGACCGAATCAGGAGTGTCTGTACAGCTTTTAACTGCAGCGGAACATGTCATTTCTCTATT 140
Db 43613 CTGGATGACCGAATCAGGAGTGTCTGTACAGCTTTTAACTGCAGCGGAACATGTCATTTCTCTATT 43672
QY 141 TCTGTGACGGCCCCCATGCCCCCACCACCAAGAGGCGAGGAGCCCGCATGATGAC 200
Db 43673 TCTGTGACGGCCCCCATGCCCCCACCACCAAGAGGCGAGGAGCCCGCATGATGAC 43732
QY 201 CCTCGTCTCTGCAGGGTGTCTGTACAGCTTTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db 43733 CCTCGTCTCTGCAGGGTGTCTGTACAGCTTTTCTGAAAGCTTTAGACAGTAACAGTTTG 43792
QY 261 CACATGAAAAATAAAGCGAGCACCTAGACAA 291
Db 43793 CACATGAAAAATAAAGCGAGCACCTAAACAA 43823

RESULT 6

US-09-949-016-12954
; Sequence 12954, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12954
; LENGTH: 46492
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(46492)
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-12954

Query Match 91.6%; Score 269.4; DB 3; Length 46492;
Best Local Similarity 99.6%; Pred. No. 9.1e-80;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 TCCCATGATGCTCTGTCAAGGAGTGGAACAATCACCTACTGAGGAGCCACCCCGCCT 80
Db 43553 TCCCATGATGCTCTGTCAAGGAGTGGAACAATCACCTACTGAGGAGCCACCCCGCCT 43612
QY 81 CTGGATGACCGAATCAGGAGTGTCTGTACAGCTTTTAACTGCAGCGGAACATGTCATTTCTCTATT 140
Db 43613 CTGGATGACCGAATCAGGAGTGTCTGTACAGCTTTTAACTGCAGCGGAACATGTCATTTCTCTATT 43672


```
Db 159 AGCAATTTCTCTGGCTCAGCCTCTAGAGTTGCTGGAGTACAGGCACCAAGCCACTGTACC 218
; LENGTH: 12128
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-12587

Qy 125 ATGTCATTCTCTATTCTGTGACGCGCCCATCGTCCCGCCACCCACCCACCAAGAGGCGAGG 184
; Mismatches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Db 219 TGGCTTCATTGTTTGTGCTGAGCATCAATCTTCCACTCCACTGTATGCTCCATA 278
; Length 12128; Score 31.4; DB 3;

Qy 185 AGGCGCCAGTCATGACCCCTGCTCTGCAGG 215
; Best Local Similarity 54.9%; Pred. No. 7.4;
; Mismatches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Db 279 AGGCGCCAGACAGCTGTCAATATATGAGG 309
; Length 12128; Score 31.4; DB 3;

RESULT 10
US-09-949-016-14722/c
; Sequence 14722, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14722
; LENGTH: 32104
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-14722

Query Match 12.2%; Score 35.8; DB 3; Length 32104;
Best Local Similarity 52.3%; Pred. No. 0.36;
Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Qy 65 AGCCCGCCAGCTCTGATGACCGAATCAGGACTAGAGTTAACTGACGCGGAAC 124
; Mismatches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Db 5051 AGCAATTTCTCTGCGCTCAGCCTCTAGAGTTGCTGGAGTACAGGCACCAAGCCACTGTACC 4992
; Length 12129; Score 31.4; DB 3;

Qy 125 ATGTCATTCTCTATTCTGTGACGCGCCCATCGTCCCGCCACCCACCAAGAGGCGAGG 184
; Best Local Similarity 54.9%; Pred. No. 7.4;
; Mismatches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Db 4991 TGGCTTCATTGTTTGTGCTGAGCATCAATCTTCCACTCCACTGTATGCTCCATA 4932
; Length 12129; Score 31.4; DB 3;

Qy 185 AGGCGCCAGTCATGACCCCTGCTCTGCAGG 215
; Best Local Similarity 54.9%; Pred. No. 7.4;
; Mismatches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Db 4931 AGGCGCCAGACAGCTGTCAATATGATGAGG 4901
; Length 12129; Score 31.4; DB 3;

RESULT 11
US-09-949-016-12587/c
; Sequence 12587, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12587
```

```
; LENGTH: 12128
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-12587

Query Match 10.7%; Score 31.4; DB 3; Length 12128;
Best Local Similarity 54.9%; Pred. No. 7.4;
Matches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 64 GAGCCCGCCAGCTCTGATGACCGAATCAGGACTAGAGTTAACTGACGCGGA 123
; Mismatches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Db 2147 GACCGCGCGCCCTCTCCGGCTCCCACTTCCAGCGGCTGAGTTCACTTTCCACACCC 2088
; Length 12129; Score 31.4; DB 3;

Qy 124 CATGTCATTCTCTATTCTGTGACGCGCCCATCGTCCCGCCACCCACCA 176
; Best Local Similarity 54.9%; Pred. No. 7.4;
; Mismatches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Db 2087 CTTTTCAGCGCCACCCCTCCCTGACCGCGCTCCCACTTCCCGTCCCGCAGCCA 2035
; Length 12129; Score 31.4; DB 3;

RESULT 12
US-09-949-016-15713/c
; Sequence 15713, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15713
; LENGTH: 12129
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-15713

Query Match 10.7%; Score 31.4; DB 3; Length 12129;
Best Local Similarity 54.9%; Pred. No. 7.4;
Matches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Qy 64 GAGCCCGCCAGCTCTGATGACCGAATCAGGACTAGAGTTAACTGACGCGGA 123
; Mismatches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Db 2147 GACCGCGCGCCCTCTCCGGCTCCCACTTCCAGCGGCTGAGTTCACTTTCCACACCC 2088
; Length 12129; Score 31.4; DB 3;

Qy 124 CATGTCATTCTCTATTCTGTGACGCGCCCATCGTCCCGCCACCA 176
; Best Local Similarity 54.9%; Pred. No. 7.4;
; Mismatches 62; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Db 2087 CTTTTCAGCGCCACCCCTCCCTGACCGCGCTCCCACTTCCCGTCCCGCAGCCA 2035
; Length 12129; Score 31.4; DB 3;

RESULT 13
US-09-489-039A-5265/c
; Sequence 5265, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5265
; LENGTH: 1809
; TYPE: DNA
```

! ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5265

Query Match 10.5%; Score 30.8; DB 3; Length 1809;
Best Local Similarity 57.1%; Pred. No. 5.5;
Matches 56; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 52 TCACCTACTGAGGAGCCCGCCGCTCTGGATGACCGAATCAGGACTAGAGTTTA 111
DB 238 TCGTTCGTATAGAACAACCGCGGAGGTGAGTCTGTGNAAGCAGGATCGCGTTCA 179
QY 112 ACTGCACGGAACATGTCATTCTCTATTCTTGTGACG 149
DB 178 GCTGCGGTCAAATGCCACTTCTCTTACCGTGGCG 141

RESULT 14
US-08-513-846-14
; Sequence 14, Application US/08513846
; Patent No. 5792609
; GENERAL INFORMATION:
; APPLICANT: WATAIYA, YUSUKE
; APPLICANT: YAMANE, AKIO
; TITLE OF INVENTION: DETECTION OF MALARIA
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/513.846
; FILING DATE: 11-SEP-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP94/00324
; FILING DATE: 28-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5/52541
; FILING DATE: 12-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 209-037-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-513-846-14

Query Match 10.4%; Score 30.6; DB 2; Length 150;
Best Local Similarity 65.2%; Pred. No. 2.4;
Matches 45; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 1 CGGCGATCGCGCGGCTTATTCCTGATGCTCTGTCAAGGAGTGGACAATCACCTACT 60
DB 73 CGGCGATCGCGCGGCTTATTCCTGATGCTCTGTCAAGGAGTGGACAATCACCTACT 132
QY 61 GAGGAGCCC 69

DB 133 TTGGGTTCC 141

RESULT 15
US-09-385-982-453
; Sequence 453, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE OF INVENTION: PRODUCTS: 11
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 453
; LENGTH: 380
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(380)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-453

Query Match 10.4%; Score 30.6; DB 3; Length 380;
Best Local Similarity 65.2%; Pred. No. 3.5;
Matches 45; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 1 CGGCGATCGCGCGGCTTATTCCTGATGCTCTGTCAAGGAGTGGACAATCACCTACT 60
DB 270 CGGCGATCGCGCGGCTTATTCCTGATGCTCTGTCAAGGAGTGGACAATCACCTACT 329
QY 61 GAGGAGCCC 69
DB 330 TTGGGTTCC 338

Search completed: December 6, 2005, 11:47:23
Job time : 130.321 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 1175.67 Seconds
(without alignments)
4099.008 Million cell updates/sec

Title: US-09-980-046B-8
Perfect score: 103
Sequence: 1 cggacagacgcagcagtc.....tgagcccttcctcacaaaa 103

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_ges1:*
10: gb_ges2:*
11: gb_ges3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	
1	103	100.0	303	2	BG181484	
C	2	101.4	98.4	180	1	AI990634
	3	101.4	98.4	302	2	BG196487
4	101.4	98.4	304	2	BG203670	
C	5	101.4	98.4	532	3	BM993129
	6	101.4	98.4	538	1	AI141684
C	7	101.4	98.4	601	3	BQ004264
	8	101.4	98.4	611	5	BX099653
C	9	101.4	98.4	670	6	CA310277
	10	101.4	98.4	672	3	BM666896
C	11	101.4	98.4	672	5	BU628665
	12	101.4	98.4	673	3	BM982024
C	13	101.4	98.4	687	6	CD742559
	14	101.4	98.4	1874	4	BC025710
C	15	100.4	97.5	641	2	BF439942
	16	100.4	97.5	690	2	BG054851
C	17	99.8	96.9	304	2	BG192962
	18	99.8	96.9	638	5	BU731752
C	19	98.4	95.5	223	1	AW071613
	20	98.4	95.5	1841	4	CR607140
C	21	98.2	95.3	304	2	BG194006
	22	98.2	95.3	421	1	AA826047

C	23	98.2	95.3	480	1	AA977801
C	24	97	94.2	493	2	BE677186
C	25	96.8	94.0	457	8	N50984
C	26	96.6	93.8	251	1	AW084330
C	27	96.6	93.8	303	2	BG214210
C	28	96.4	93.6	321	1	AA252172
C	29	96.4	93.6	500	3	BP392226
C	30	95	92.2	244	1	AA931748
C	31	90.4	87.8	268	1	AI280735
C	32	90.4	87.8	419	2	BG686255
C	33	87.8	85.2	332	1	AA994205
C	34	86	83.5	622	3	BM999231
C	35	83.2	80.8	374	8	T64144
C	36	75.4	73.2	300	1	AI361098
C	37	75.4	73.2	417	1	AW014432
C	38	75.4	73.2	701	6	CD366471
C	39	74.4	72.2	295	1	AW517751
C	40	74.4	72.2	322	1	AW238793
C	41	74.4	72.2	345	3	BM126146
C	42	74.4	72.2	379	3	BM126439
C	43	74.4	72.2	484	3	BM679324
C	44	74.4	72.2	690	2	BF969846
C	45	74.4	72.2	706	3	BM679069

ALIGNMENTS

RESULT 1
BG181484
LOCUS
DEFINITION
RST332 Athersys RAGE Library Homo sapiens cdNA, mRNA sequence.
ACCESSION
BG181484
VERSION
BG181484.1 GI:13703171
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 303)
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
Offenbacher,J., Danzig,J. and Ducar,M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
11329013
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 303.
Location/Qualifiers
1. 303
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

JOURNAL
PUBMED
COMMENT

FEATURES
source

ORIGIN

Query Match 100.0%; Score 103; DB 2; Length 303;

Best Local Similarity 100.0%; Pred. No. 2.4e-23;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACAGCGGCGAGCTCCAGCTCTGGTTCTCTCGGTTTATCTGTAGATGAA 60
|||||
Db 193 CGGACAGCGGCGAGCTCCAGCTCTGGTTCTCTCGGTTTATCTGTAGATGAA 252
|||||
QY 61 ATGTTTCCCAATAAATAGGGCGCATGAGCCCTTCTCACAATAA 103
|||||
Db 253 ATGTTTCCCAATAAATAGGGCGCATGAGCCCTTCTCACAATAA 295
|||||

RESULT 2
LOCUS AI990634/c 180 bp mRNA linear EST 08-SEP-1999
DEFINITION W222408.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2497959 3',
mRNA sequence.
ACCESSION AI990634
VERSION AI990634.1 GI:5837515
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 180)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-t@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.lnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco.

FEATURES
source
1..180
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE:2497959"
/tissue_type="pooled germ cell tumors"
/lab_hosts="DH10B"
/clone_lib="NCI_CGAP_GC6"

/notes="Vector: p773D-Fac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA
from the normalized library NCI_CGAP GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (Cloneds
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 98.4%; Score 101.4; DB 1; Length 180;
Best Local Similarity 99.0%; Pred. No. 7.3e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGCGGCGAGCTCCAGCTCTGGTTCTCTCGGTTTATCTGTAGATGAA 60
|||||
Db 108 CGGACAGCGGCGAGCTCCAGCTCTGGTTCTCTCGGTTTATCTGTAGATGAA 49
|||||
QY 61 ATGTTTCCCAATAAATAGGGCGCATGAGCCCTTCTCACAATAA 103
|||||
Db 48 ATGTTTCCCAATAAATAGGGCGCATGAGCCCTTCTCACAATAA 6

RESULT 3
LOCUS BG196487
DEFINITION RST15710 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG196487
VERSION BG196487.1 GI:13718174
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 302)

Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,
Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,
Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K.,
Offenbacher, J., Danzig, J., and Ducar, M.

Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
11329013

CONTACT: Scott J. Cain

Athersys, Inc.

3201 Carnegie Ave., Cleveland, OH 44115, USA

Tel: 216 431 9900

Fax: 216 361 9596

Email: scain@athersys.com

High quality sequence stop: 302.

Location/Qualifiers

1..302

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/cell_line="HT1080"

/clone_lib="Athersys RAGE Library"

/notes="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression', the
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 98.4%; Score 101.4; DB 2; Length 302;
Best Local Similarity 99.0%; Pred. No. 8.1e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGCGGCGAGCTCCAGCTCTGGTTCTCTCGGTTTATCTGTAGATGAA 60
|||||
Db 191 CGGACAGCGGCGAGCTCCAGCTCTGGTTCTCTCGGTTTATCTGTAGATGAA 250
|||||
QY 61 ATGTTTCCCAATAAATAGGGCGCATGAGCCCTTCTCACAATAA 103
|||||
Db 251 ATGTTTCCCAATAAATAGGGCGCATGAGCCCTTCTCACAATAA 293
|||||

RESULT 4

LOCUS BG203670

DEFINITION RST23058 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG203670

VERSION BG203670.1 GI:13725357

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 304)

Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,

Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K., Offenbacher, J., Danzig, J., and Ducar, M.
 Creation of genome-wide protein expression libraries using random activation of gene expression
 Nat. Biotechnol. 19 (5), 440-445 (2001)
 11329013
 Contact: Scott J. Cain
 Athersys, Inc. Cleveland, OH 44115, USA
 3201 Carnegie Ave., Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596
 Email: scain@atersys.com
 High quality sequence stop: 304.
 Location/Qualifiers

FEATURES

source

1. .304

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/cell_line="HT1080"

/clone_lib="Athersys RAGE Library"

/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 98.4%; Score 101.4; DB 2; Length 304;
 Best Local Similarity 99.0%; Pred. No. 8.1e-23;
 Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTCTCTCGTTTATCTGTAGATGAA 60
 Db 193 CGGACAGCGCAGCAGTCCAGCTCTGGTTCTCTCGTTTATCTGTAGATGAA 252
 Qy 61 ATGGTCCCAATAATAGGGCATGAGCCCTCTCTCAAAAA 103
 Db 253 ATGGTCCCAATAATAGGGCATGAGCCCTCTCTCAAAAA 295

RESULT 5
 BM993129/c
 LOCUS
 DEFINITION
 IMAGE:5866052 3', mRNA sequence.
 ACCESSION
 BM993129
 VERSION
 BM993129.1 GI:19712518
 KEYWORDS
 EST.
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 532)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Jose Mercuende
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source

1. .532

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5866052"
 /tissue_type="Metastatic Chondrosarcoma"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP DT0"
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: Ecor I; Site 2: Not I;
 NCI CGAP DT0 is a cDNA library containing the following tissue(s): Metastatic Chondrosarcoma in Lung. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is
 AACTGTCGG.
 TAG TISSUE=lung metastatic chondrosarcoma
 TAG LIB=UI-H-DT0
 TAG_SEQ=AACTGTCGG"

ORIGIN

Query Match 98.4%; Score 101.4; DB 3; Length 532;
 Best Local Similarity 99.0%; Pred. No. 9.1e-23;
 Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTCTCTCGTTTATCTGTAGATGAA 60
 Db 114 CGGACAGCGCAGCAGTCCAGCTCTGGTTCTCTCGTTTATCTGTAGATGAA 55
 Qy 61 ATGGTCCCAATAATAGGGCATGAGCCCTCTCTCAAAAA 103
 Db 54 ATGGTCCCAATAATAGGGCATGAGCCCTCTCTCAAAAA 12

RESULT 6

AI141684/c

LOCUS

DEFINITION

ot08405.x1 NCI CGAP GC3 Homo sapiens cDNA clone IMAGE:1614249 3', similar to TR:009014 009014 PEPTIDE/HISTIDINE TRANSPORTER. ;, mRNA

sequence.

AI141684

VERSION

AI141684.1 GI:3649141

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 538)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone Distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40ml3 fwd. EF from Amersham

High quality sequence stop: 456.

Location/Qualifiers

1. .538

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1614249"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP GC3"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector library. Library is not normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
Query Match 98.4%; Score 101.4; DB 1; Length 538;
Best Local Similarity 99.0%; Pred. No. 9.1e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACACAGCGGACGAGTCCAGCTCTGGTTTCCTTCCTCGGTTTATTCGTTAGATGAA 60
DB 113 CGACACAGCGGACGAGTCCAGCTCTGGTTTCCTTCCTCGGTTTATTCGTTAGATGAA 54

QY 61 ATGTTCCCATTAATAAGGGCGATGAGCCCTTCCTCAGAAAAA 103
DB 53 ATGTTCCCATTAATAAGGGCGATGAGCCCTTCCTCAGAAAAA 11

RESULT 7
BQ004264/c
LOCUS
DEFINITION
UI-H-E10-ayn-o-17-0-UI.s1 NCI_CGAP_E10 Homo sapiens cDNA clone
IMAGE:5841256 3', mRNA sequence.
BQ004264
BQ004264.1 GI:19729164
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 601)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
PolyA=yes.
Location/Qualifiers
1. .601
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5841256"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP E10"
/notes="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia)
with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI_CGAP E10 is a cDNA library containing the following
tissue(s): Chondrosarcoma. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1614249"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP GC3"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector library. Library is not normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
Query Match 98.4%; Score 101.4; DB 3; Length 601;
Best Local Similarity 99.0%; Pred. No. 9.3e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACACAGCGGACGAGTCCAGCTCTGGTTTCCTTCCTCGGTTTATTCGTTAGATGAA 60
DB 114 CGACACAGCGGACGAGTCCAGCTCTGGTTTCCTTCCTCGGTTTATTCGTTAGATGAA 55

QY 61 ATGTTCCCATTAATAAGGGCGATGAGCCCTTCCTCAGAAAAA 103
DB 54 ATGTTCCCATTAATAAGGGCGATGAGCCCTTCCTCAGAAAAA 12

RESULT 8
BX099653
LOCUS
DEFINITION
611 bp mRNA linear EST 06-FEB-2003
BX099653 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:998115131 ; IMAGE:127958, mRNA sequence.
BX099653
BX099653.1 GI:27830132
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 611)
Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Radelof, U., Schneider, D. and Korn, B.
Human Unigeneset - RZPD3
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGp998115131.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
http://www.rzpd.de/CloneCards/cgi-
bin/showLib.pl.cgi/responsefileNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACAGGAACAGCTATGAC.
Location/Qualifiers
1. .611
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGp998115131 ; IMAGE:127958"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/notes="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer

```

[5', AACTGGAGAGATTAATAAGATCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 98.4%; Score 101.4; DB 5; Length 611;
 Best Local Similarity 99.0%; Pred. No. 9.3e-23;
 Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTTATTCTGTAGATGAA 60
 |||
 Db 501 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTTATTCTGTAGATGAA 560
 |||

Qy 61 ATGGTTCCTAATAAAGGGCGATGAGCCCTTCCTCAAAAA 103
 |||
 Db 561 ATGGTTCCTAATAAAGGGCGATGAGCCCTTCCTCAAAAA 603
 |||

RESULT 9

CA310277/c
 LOCUS
 DEFINITION UI-H-Ftl-big-h-15-0-UI.s1 NCI CGAP Ftl Homo sapiens cDNA clone
 UI-H-Ftl-big-h-15-0-UI 3', mRNA sequence.

ACCESSION CA310277 GI:24473331

VERSION EST.

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE

1 (bases 1 to 670)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Gary W. Hunninghake, U of I
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source

Location/Qualifiers
 1..670
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-Ftl-big-h-15-0-UI"
 /tissue_type="Alveolar Macrophage"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP Ftl"
 /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
 NCI CGAP Ftl is a normalized cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The mRNA samples were a mixture of these conditions (times refer to incubations following isolation by bronchoalveolar lavage) (some normal donor macrophages were cultured in some of the conditions, other donor macrophages in different conditions). The mRNA samples were pooled for library construction. Control 0 hours; control 3 hours; control 24 hours; LPS 100 ng/ml, 3 hours; LPS 100 ng/ml, 24 hours; PMA 10 ng/ml, 3 hours; PMA 10 ng/ml, 24 hours; Staph aureus moi 10, 3 hours; Klebsiella moi 10, 24 hours; Staph aureus moi 10, 3 hours; Staph aureus moi 10, 24 hours; Adenoviral vector (Ad5 CMV eGFP), moi 500, 3 hours; Adenoviral vector (Ad5 CMV eGFP), moi 500, 24 hours; wt adenovirus moi 500, 3 hours; wt adenovirus moi 500, 24 hours; Ad vector + LPS 3 hours; Ad vector + LPS 24 hours; wt adenovirus + LPS 3 hours; wt adenovirus + LPS 24 hours. The library was normalized according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GGCATGCGC. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.
 TAG_TISSUE=Human Lung Alveolar Macrophage
 TAG_LIB=UI-H-Ftl
 TAG_SEQ=GGCATGCGC"

ORIGIN

Query Match 98.4%; Score 101.4; DB 6; Length 670;
 Best Local Similarity 99.0%; Pred. No. 9.5e-23;
 Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTTATTCTGTAGATGAA 60
 |||
 Db 114 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTTATTCTGTAGATGAA 55
 |||

Qy 61 ATGGTTCCTAATAAAGGGCGATGAGCCCTTCCTCAAAAA 103
 |||
 Db 54 ATGGTTCCTAATAAAGGGCGATGAGCCCTTCCTCAAAAA 12
 |||

RESULT 10

BM666896/c

LOCUS

DEFINITION UI-E-CL1-aes-j-23-0-UI.s1 UI-E-CL1 Homo sapiens cDNA clone
 UI-E-CL1-aes-j-23-0-UI 3', mRNA sequence.

ACCESSION BM666896

VERSION EST.

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 672)
 Bonaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene discovery
 Genome Res. 6 (9), 791-806 (1996)
 8889548

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRP, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
 Seq primer: M13 Forward
 POLYA=Yes

FEATURES

source

Location/Qualifiers
 1..672
 /organism="Homo sapiens"
 /mol_type="mRNA"

```

/db xref="taxon:9606"
/clone="UI-E-CL1-aes-j-23-0-UI"
/tissue_type="human retina"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CL1"
/notes="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-CL1 is a normalized cDNA library containing the
following tissue(s): retina. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CCGCG. This library was
created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI).
TAG_TISSUE=human retina
TAG_LIB=UI-E-CL1
TAG_SEQ=CCGCG"

```

```

ORIGIN
Query Match      98.4%; Score 101.4; DB 3; Length 672;
Best Local Similarity 99.0%; Pred. No. 9.5e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATCTGTAGATGAA 60
    |||
Db 117 CGGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATCTGTAGATGAA 58
    |||

QY 61 ATGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 103
    |||
Db 57 ATGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 15
    |||

```

```

RESULT 11
BU628665/c
LOCUS
DEFINITION
UI-H-PG0-bdf-1-04-0-UI.s1 NCI CGAP_EN1_2 Homo sapiens cDNA clone
UI-H-PG0-bdf-1-04-0-UI 3', mRNA sequence.
ACCESSION
BU628665
VERSION
BU628665.1 GI:23294879
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 672)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapb@mail.nih.gov
Tissue Procurement: James Martin
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone Distribution Information can be obtained
from Dr. M. Bento Soares, bent-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1. .672
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-PG0-bdf-1-04-0-UI"

```

```

FEATURES
source
1. .672
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-PG0-bdf-1-04-0-UI"

```

```

/tissue_type="Enchondroma cell line"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP_EN1_2"
/notes="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP_EN1_2 is a cDNA library containing the following
tissue(s): Enchondroma cell line (2 cell lines). The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is CCGGTCACTC. The cell lines was provided by Dr
James Martin from University of Iowa.
TAG_TISSUE=Enchondroma cell line (Mix of EN1 and EN2)
TAG_LIB=UI-H-FG0
TAG_SEQ=CCGTCACTC"

```

```

ORIGIN
Query Match      98.4%; Score 101.4; DB 5; Length 672;
Best Local Similarity 99.0%; Pred. No. 9.5e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATCTGTAGATGAA 60
    |||
Db 114 CGGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATCTGTAGATGAA 55
    |||

QY 61 ATGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 103
    |||
Db 54 ATGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 12
    |||

```

```

RESULT 12
BM982024/c
LOCUS
DEFINITION
UI-CF-EN1-adg-j-16-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
UI-CF-EN1-adg-j-16-0-UI 3', mRNA sequence.
ACCESSION
BM982024
VERSION
BM982024.1 GI:19605106
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 673)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
AUTHORS
Normalization and subtraction: two approaches to facilitate gene
discovery
TITLE
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
8889548
COMMENT
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers

```

```

FEATURES

```



```

source
1. .673
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EN1-adg-j-16-0-UI"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-CF-EN1"
/notes="Organ: Lung; Vector: p7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (d)18 tail. The sequence tag for this library is CTGCTCAGGT. TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS 6hr to LPS 24h TAG LIB=UI-CF-EN1 TAG_SEQ=CTGCTCAGGT"

ORIGIN
Query Match 98.4%; Score 101.4; DB 3; Length 673;
Best Local Similarity 99.0%; Pred. No. 9.5e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCCTCGGTTTATTCGTAGATGAA 60
Db 117 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCCTCGGTTTATTCGTAGATGAA 58

Qy 61 ATGGTTCCCATAAATAGGGCATGAGCCCTTCCTCAAAAA 103
|||||
Db 57 ATGGTTCCCATAAATAGGGCATGAGCCCTTCCTCAAGAAA 15

RESULT 13
CD742559/c 687 bp mRNA linear EST 05-AUG-2004
LOCUS UI-H-FT2-bj1-e-08-0-UI.s1 NCI CGAP FT2 Homo sapiens cDNA clone
DEFINITION UI-H-FT2-bj1-e-08-0-UI 3', mRNA sequence.
ACCESSION CD742559
VERSION CD742559.1 GI:32293409
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 687)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NCI
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Gary W. Hunninghake, U of I
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/cgap.html
Seq primer: M13 FORWARD

FEATURES
source
1. .687
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FT2-bj1-e-08-0-UI"
/tissue_type="Alveolar Macrophage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FT2"
/notes="Organ: Lung; Vector: p7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FT2 is a subtracted cDNA library constructed from a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The mRNA samples were a mixture of these conditions (times refer to incubations following isolation by bronchoalveolar lavage) (some normal donor macrophages were cultured in some of the conditions, other donor macrophages in different conditions). The mRNA samples were pooled for library construction. Control 0 hours; control 3 hours; control 24 hours; LPS 100 ng/ml, 3 hours; LPS 100 ng/ml, 24 hours; PMA 10 ng/ml, 3 hours; PMA 10 ng/ml, 24 hours; Klebsiella aureus moi 10, 3 hours; Staph aureus moi 10, 24 hours; Adenoviral vector (Ad5 CMV eGFP), moi 500, 3 hours; Adenoviral vector (Ad5 CMV eGFP), moi 500, 24 hours; wt adenovirus moi 500, 3 hours; wt adenovirus moi 500, 24 hours; Ad vector + LPS 3 hours; Ad vector + LPS 24 hours; wt adenovirus + LPS 3 hours; wt adenovirus + LPS 24 hours. The library was subtracted according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa. TAG TISSUE=Human Lung Alveolar Macrophage TAG LIB=UI-H-FT2 TAG_SEQ=GGCCATGCGG"

ORIGIN
Query Match 98.4%; Score 101.4; DB 6; Length 687;
Best Local Similarity 99.0%; Pred. No. 9.5e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCCTCGGTTTATTCGTAGATGAA 60
|||||
Db 111 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCCTCGGTTTATTCGTAGATGAA 52

Qy 61 ATGGTTCCCATAAATAGGGCATGAGCCCTTCCTCAAAAA 103
|||||
Db 51 ATGGTTCCCATAAATAGGGCATGAGCCCTTCCTCAAAAA 9

RESULT 14
BC025710 1874 bp mRNA linear HTC 19-NOV-2003
LOCUS BC025710
DEFINITION Homo sapiens solute carrier family 15, member 3, mRNA (cDNA clone IMAGE5224878), containing frame-shift errors.
ACCESSION BC025710
VERSION BC025710.1 GI:19343588
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1874)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heise, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
```

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Vallalao, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, J.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932
2 (bases 1 to 1874)
Strausberg, R.
Direct Submission
Submitted (06-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NTH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nri.nih.gov
Akhter, N., Ayte, K., Becketrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.J., Pearson, C., Maskeri, B., Masrrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 49 Row: m Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7706116
This clone has the following problem: frame shifted.

FEATURES
source
1..1874
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5224878"
/tissue_type="Pancreas, Spleen, adult pooled"
/clone_lib="NIH_MGC_120"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN
Query Match 98.4%; Score 101.4; DB 4; Length 1874;
Best Local Similarity 99.0%; Pred. No. 1.2e-21;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGACAGCGGCGAGCTCCAGCTCTGGTTTCCTTCGGTTTATCTGTTAGATGAA 60
DB 1707 CGGACAGCGGCGAGCTCCAGCTCTGGTTTCCTTCGGTTTATCTGTTAGATGAA 1766
QY 61 ATGTTTCCCATATAAAGGGCGATGAGCCCTTCTCACAAAA 103
DB 1767 ATGTTTCCCATATAAAGGGCGATGAGCCCTTCTCACAAAA 1809

RESULT 15

BF439942/c

LOCUS

DEFINITION

BF439942 641 bp mRNA linear EST 29-NOV-2000
nacsia09.x1 NCI CGAP Brn23 Homo sapiens cDNA clone IMAGE:3405976 3' similar to TR:Q9WU80 Q9WU80 CAMP INDUCIBLE 1 PROTEIN. ;, mRNA sequence.

ACCESSION

BF439942

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 641)
NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGP), Tumor Gene Index
Unpublished (1998)
Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 450.

FEATURES

source

1..641

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3405976"

/tissue_type="glioblastoma (pooled)"

/lab_host="DH10B"

/clone_lib="NCI CGAP Brn23"

/note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo (dT) primer [5' TGTACCAATCTGAAGTGGAGCGCGCATATCTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match

Best Local Similarity

Matches 101; Conservative

0; Mismatches

1; Indels

0; Gaps

0;
QY 1 CGGACAGCGGCGAGCTCCAGCTCTGGTTTCCTTCGGTTTATCTGTTAGATGAA 60
DB 102 CGGACAGCGGCGAGCTCCAGCTCTGGTTTCCTTCGGTTTATCTGTTAGATGAA 43
QY 61 ATGTTTCCCATATAAAGGGCGATGAGCCCTTCTCACAAAA 102
DB 42 ATGTTTCCCATATAAAGGGCGATGAGCCCTTCTCACGAAA 1

Search completed: December 7, 2005, 09:56:00

Job time : 1178.77 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 265.316 Seconds
(without alignments)
6427.437 Million cell updates/sec

Title: US-09-980-046B-9

Perfect score: 30

Sequence: 1 cggaaataaaggctgtgtaagacacaaaa 30

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_ro.*

10: gb_sts.*

11: gb_sy.*

12: gb_un.*

13: gb_vl.*

14: gb_btg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28.4	94.7	619	BC009507	BC009507 Homo sapi
2	28.4	94.7	668	BD203732	BD203732 Human nuc
3	28.4	94.7	668	AX014887	AX014887 Sequence
4	28.4	94.7	670	BD134423	BD134423 Human nuc
5	28.4	94.7	670	AX017252	AX017252 Sequence
6	28.4	94.7	670	AX524956	AX524956 Sequence
7	26.4	88.0	163399	AC079860	AC079860 Homo sapi
8	26.4	88.0	186759	AL645608	AL645608 Human DNA
9	26.4	88.0	200491	AB107102	AB107102 Homo sapi
10	26.4	88.0	327883	BX571684	BX571684 Homo sapi
11	24.4	81.3	509	CQ922970	CQ922970 Sequence
12	23.4	78.0	161151	AP000876	AP000876 Homo sapi
13	23.4	78.0	211945	AC006238	AC006238 Homo sapi
14	23	76.7	301	BD275867	BD275867 COMPOUNDS
15	23	76.7	301	AR220652	AR220652 Sequence
16	23	76.7	301	AR255646	AR255646 Sequence
17	23	76.7	301	AR281216	AR281216 Sequence
18	23	76.7	301	AR541237	AR541237 Sequence

C	19	23	76.7	301	6	AX365911	Sequence
	20	22.6	75.3	102976	15	AC142224	Medicago
	21	22.6	75.3	199961	9	AL596111	Mouse DNA
	22	22.2	74.0	194232	14	AC154523	Mus muscu
C	23	22.2	74.0	208143	14	AC164357	Bos tauru
C	24	22.2	74.0	236536	9	AC139941	Mus muscu
	25	22	73.3	447	9	BC062266	Mus muscu
	26	22	73.3	140864	9	AC126803	Mus muscu
C	27	22	73.3	141855	5	CR388381	zebrafish
	28	22	73.3	154090	14	AC142024	Rattus no
C	29	22	73.3	169547	9	AL590503	Mouse DNA
	30	22	73.3	178124	9	AC153970	Mus muscu
C	31	22	73.3	179919	14	AC114383	Rattus no
C	32	22	73.3	181034	14	AC135486	Rattus no
	33	22	73.3	187002	5	CR391941	zebrafish
C	34	22	73.3	206168	9	AC153912	Mus muscu
	35	22	73.3	223809	14	AC105474	Rattus no
C	36	22	73.3	238611	14	AC108953	Rattus no
C	37	22	73.3	244255	14	AC109041	Rattus no
C	38	22	73.3	246384	14	AC108543	Rattus no
	39	22	73.3	265663	14	AC109741	Rattus no
C	40	22	73.3	331363	14	AC113829	Rattus no
C	41	21.8	72.7	189793	14	AC149188	Papio anu
C	42	21.6	72.0	17292	8	HSBPB72EX	H.sapiens e
	43	21.6	72.0	19269	8	AL161784	Human DNA
	44	21.6	72.0	98274	8	HS714B7	Human DNA s
C	45	21.6	72.0	105412	14	AP007403	Lotus cor

ALIGNMENTS

RESULT 1	BC009507	Homo sapiens interferon, alpha-inducible protein (clone IFI-15K), mRNA (cdna clone MGC:3945 IMAGE:3545944), complete cds.	619 bp	mena	linear	PRI 08-MAR-2005
LOCUS	BC009507					
DEFINITION	BC009507					
ACCESSION	BC009507					
VERSION	BC009507.2	GI:38114701				
KEYWORDS	MSC.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 619)					
AUTHORS	Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Brownstein MJ, Usdin TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullaly SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH, Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW, Villalón DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Kertanen M, Madan A, Rodriguez S, Sanchez A, Whiting M, Madan A, Young AC, Shevchenko Y, Bouffard GG, Blackesley RM, Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield JS, Krzywinski MI, Skalska U, Smalusz J, Schnerch A, Schein JF, Jones SJ and Marra MA.					
CONSTRM	Mammalian Gene Collection Program Team					
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences					
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)					
PUBMED	12477932					
REFERENCE	2 (bases 1 to 619)					
AUTHORS	Director MGC Project.					
TITLE	Direct Submission					
JOURNAL	Submitted (25-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA					


```
BD134423
LOCUS       BD134423               670 bp      DNA      linear      PAT 18-SEP-2002
DEFINITION   Human nucleic acid sequence originating in mammary tumor tissue.
ACCESSION    BD134423
VERSION      BD134423.1  GI:23229368
KEYWORDS     JP 2002506643-A/1.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Homnidae; Homo.
REFERENCE    1 (bases 1 to 670)
AUTHORS      Speft,T., Hintzman,B., Armin,S., Pirarski,C., Edgar,D. and
             Rosenthal,A.
TITLE        Human nucleic acid sequence originating in mammary tumor tissue
JOURNAL      Patent: JP 2002506643-A 1 05-MAR-2002;
             METAGEN GESELLSCHAFT FUER GENOME FORSCHUNG MBH
COMMENT      OS Homo sapiens (human)
             PN JP 2002506643-A/1
             PD 05-MAR-2002
             PF 19-MAR-1999 JP 2000536852
             PR 20-MAR-1998 DE 198 13 839.3
             PI THOMAS SPEFT, BERND HINTZMAN, SHCMITT ARMIN, CHRISTIAN PIRARSKI,
             DUHL EDGAR,
             PI ANDRE ROSENTHAL
             PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P35/00,C07K14/47,
             C07K16/18,
             PC C12N15/10,C12N15/00,A61K37/02,C12N5/00 CC Human
             PC C12N1/19,C12N5/10,C12N15/00,A61K37/02,C12N5/00 CC
             nucleic acid sequence originating in mammary tumor
             tissue
FEATURES     source
             FH Key          Location/Qualifiers
             FT source       1..670
                             /organism="Homo sapiens (human)".
                             Location/Qualifiers
                             1..670
                             /organism="Homo sapiens"
                             /mol_type="genomic DNA"
                             /db_xref="taxon:9606"

ORIGIN
Query Match      94.7%; Score 28.4; DB 6; Length 670;
Best Local Similarity 96.7%; Pred. No. 5;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGGAATAAAGCGCTGTTGTTAAAGACAAAAA 30
    |||||
Db 637 CGGAATAAAGCGCTGTTGTTAAAGACAAAAA 666

RESULT 5
AC079860/c
LOCUS       AC079860               670 bp      DNA      linear      PAT 07-SEP-2000
DEFINITION   Sequence 2 from Patent WO9947669.
ACCESSION    AC079860
VERSION      AC079860.1  GI:10042170
KEYWORDS     AX017252
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Homnidae; Homo.
REFERENCE    1
AUTHORS      Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
             Pilarsky,C.
TITLE        Human nucleic acid sequences from tissue of breast tumors
JOURNAL      Patent: WO 9947669-A 2 23-SEP-1999;
             SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
             BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
             (DE); PILARSKY CHRISTIAN (DE)
FEATURES     source
             1..670
                             /organism="Homo sapiens"
                             /mol_type="unassigned DNA"

ORIGIN
Query Match      94.7%; Score 28.4; DB 6; Length 670;
Best Local Similarity 96.7%; Pred. No. 5;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGGAATAAAGCGCTGTTGTTAAAGACAAAAA 30
    |||||
Db 637 CGGAATAAAGCGCTGTTGTTAAAGACAAAAA 666

RESULT 7
AC079860/c
LOCUS       AC079860               163399 bp   DNA      linear      HTG 23-DEC-2000
DEFINITION   Homo sapiens chromosome RPCI-11 clone RP11-671C15, WORKING DRAFT
             SEQUENCE, 54 unordereded pieces.
ACCESSION    AC079860
VERSION      AC079860.3  GI:11990750
KEYWORDS     HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Homnidae; Homo.
REFERENCE    1 (bases 1 to 163399)
AUTHORS      Waterston,R.H.
TITLE        The sequence of Homo sapiens clone
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 163399)
AUTHORS      Waterston,R.H.
TITLE        Direct Submission
JOURNAL      Submitted (14-SEP-2000) Genome Sequencing Center, Washington
             University School of Medicine, 4444 Forest Park Parkway, St. Louis,
             MO 63108, USA
COMMENT      On Dec 23, 2000 this sequence version replaced gi:10518428.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
```

```
ORIGIN
Query Match      94.7%; Score 28.4; DB 6; Length 670;
Best Local Similarity 96.7%; Pred. No. 5;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGGAATAAAGCGCTGTTGTTAAAGACAAAAA 30
    |||||
Db 637 CGGAATAAAGCGCTGTTGTTAAAGACAAAAA 666

RESULT 6
AX524956
LOCUS       AX524956               670 bp      DNA      linear      PAT 21-NOV-2002
DEFINITION   Sequence 2 from Patent EPI236799.
ACCESSION    AX524956
VERSION      AX524956.1  GI:25170038
KEYWORDS     .
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Homnidae; Homo.
REFERENCE    1
AUTHORS      Specht,T., Hinzmann,B., Schmitt,A., Pilarsky,C., Dahl,E. and
             Rosenthal,A.
TITLE        Human nucleic acid sequences derived from breast tumor tissue
JOURNAL      Patent: EP 1236799-A 2 04-SEP-2002;
             metagen Pharmaceuticals GmbH (DE)
FEATURES     source
             1..670
                             /organism="Homo sapiens"
                             /mol_type="unassigned DNA"
                             /db_xref="taxon:9606"

ORIGIN
Query Match      94.7%; Score 28.4; DB 6; Length 670;
Best Local Similarity 96.7%; Pred. No. 5;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGGAATAAAGCGCTGTTGTTAAAGACAAAAA 30
    |||||
Db 637 CGGAATAAAGCGCTGTTGTTAAAGACAAAAA 666

RESULT 7
AC079860/c
LOCUS       AC079860               163399 bp   DNA      linear      HTG 23-DEC-2000
DEFINITION   Homo sapiens chromosome RPCI-11 clone RP11-671C15, WORKING DRAFT
             SEQUENCE, 54 unordereded pieces.
ACCESSION    AC079860
VERSION      AC079860.3  GI:11990750
KEYWORDS     HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Homnidae; Homo.
REFERENCE    1 (bases 1 to 163399)
AUTHORS      Waterston,R.H.
TITLE        The sequence of Homo sapiens clone
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 163399)
AUTHORS      Waterston,R.H.
TITLE        Direct Submission
JOURNAL      Submitted (14-SEP-2000) Genome Sequencing Center, Washington
             University School of Medicine, 4444 Forest Park Parkway, St. Louis,
             MO 63108, USA
COMMENT      On Dec 23, 2000 this sequence version replaced gi:10518428.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
```

Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 ----- Project Information -----
 Center project name: H1NH0571C15
 ----- Summary Statistics -----
 Sequencing vector: M13; 100%
 Sequencing vector: plasmid; 0%
 Chemistry: Dye-terminator Big Dye; 0% of reads
 Chemistry: Dye-terminator Big Dye; 0% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 123674 bases at least Q40
 Consensus quality: 137441 bases at least Q30
 Consensus quality: 143762 bases at least Q20
 Insert size: 176000; agarose-fp
 Insert size: 158099; sum-of-contigs
 Quality coverage: 2.94 in Q20 bases; agarose-fp
 Quality coverage: 3.43 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 54 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1229: contig of 1229 bp in length
 1230: gap of unknown length
 1231: contig of 1115 bp in length
 2444: gap of unknown length
 2445: contig of 1637 bp in length
 4181: contig of 1637 bp in length
 4182: gap of unknown length
 4282: contig of 1297 bp in length
 5579: gap of unknown length
 5678: contig of 1052 bp in length
 6731: contig of 1306 bp in length
 6831: contig of 1306 bp in length
 8137: gap of unknown length
 8237: contig of 1303 bp in length
 9540: gap of unknown length
 9640: contig of 1606 bp in length
 11246: gap of unknown length
 11346: contig of 1403 bp in length
 12749: gap of unknown length
 12849: contig of 1546 bp in length
 14395: gap of unknown length
 14495: contig of 1401 bp in length
 15895: gap of unknown length
 15995: contig of 1088 bp in length
 17084: gap of unknown length
 17184: contig of 1230 bp in length
 18414: gap of unknown length
 20492: contig of 1979 bp in length
 20493: gap of unknown length
 22178: contig of 1586 bp in length
 22179: gap of unknown length
 22279: contig of 1851 bp in length
 24130: gap of unknown length
 24230: contig of 2326 bp in length
 26556: gap of unknown length
 26656: contig of 2345 bp in length
 29100: gap of unknown length
 29101: contig of 1900 bp in length
 31000: gap of unknown length
 31001: contig of 1578 bp in length
 32678: gap of unknown length
 32679: contig of 1435 bp in length
 34213: gap of unknown length
 34214: contig of 3205 bp in length
 34314: gap of unknown length
 37618: gap of unknown length
 40183: contig of 2565 bp in length
 40184: gap of unknown length
 40283: contig of 1545 bp in length
 40284: contig of 1545 bp in length

41829: gap of unknown length
 41929: contig of 3035 bp in length
 44963: gap of unknown length
 45063: contig of 3279 bp in length
 48344: gap of unknown length
 48442: contig of 1666 bp in length
 48443: gap of unknown length
 50109: gap of unknown length
 50209: contig of 2322 bp in length
 52530: gap of unknown length
 52531: contig of 2331 bp in length
 52631: gap of unknown length
 54961: contig of 2490 bp in length
 55062: gap of unknown length
 57551: contig of 2256 bp in length
 57552: gap of unknown length
 59907: contig of 2256 bp in length
 60007: gap of unknown length
 62118: contig of 2111 bp in length
 62218: gap of unknown length
 66434: contig of 4216 bp in length
 66534: gap of unknown length
 69895: contig of 3361 bp in length
 69995: gap of unknown length
 73349: contig of 3354 bp in length
 73449: gap of unknown length
 76265: contig of 2816 bp in length
 76365: gap of unknown length
 79693: contig of 3328 bp in length
 79793: gap of unknown length
 82725: contig of 2931 bp in length
 82825: gap of unknown length
 85709: contig of 2884 bp in length
 85809: gap of unknown length
 89624: contig of 3816 bp in length
 89724: gap of unknown length
 93733: contig of 4009 bp in length
 93833: gap of unknown length
 97156: contig of 3323 bp in length
 97257: gap of unknown length
 100918: contig of 3662 bp in length
 101019: gap of unknown length
 101019: contig of 3956 bp in length
 105074: gap of unknown length
 105075: contig of 5748 bp in length
 110822: gap of unknown length
 110823: contig of 4714 bp in length
 115637: gap of unknown length
 115737: contig of 4740 bp in length
 120477: gap of unknown length
 120577: contig of 5098 bp in length
 125675: gap of unknown length
 125775: contig of 6020 bp in length
 131795: gap of unknown length
 131895: contig of 6224 bp in length
 138119: gap of unknown length
 138219: contig of 5412 bp in length
 143631: gap of unknown length
 143731: contig of 5024 bp in length
 148755: gap of unknown length
 148855: contig of 7245 bp in length
 156100: gap of unknown length
 156200: contig of 7200 bp in length.

FEATURES
 source
 1..163399
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosomes="RPC1-11"
 /clone="Rp11-671C15"
 misc_feature
 1..1229
 /notes="assembly_name:Contig22"
 1230..1329
 /estimated_length=unknown
 1330..2444
 /notes="assembly_name:Contig25"

```

gap          2445. .2544
             /estimated_length=unknown
misc_feature 2545. .4181
             /note="assembly_name:Contig29"
gap          4182. .4281
             /estimated_length=unknown
misc_feature 4282. .5578
             /note="assembly_name:Contig30"
gap          5579. .5678
             /estimated_length=unknown
misc_feature 5679. .6730
             /note="assembly_name:Contig31"
gap          6731. .6830
             /estimated_length=unknown
misc_feature 6831. .8136
             /note="assembly_name:Contig32"
gap          8137. .8236
             /estimated_length=unknown
misc_feature 8237. .9539
             /note="assembly_name:Contig33"
gap          9540. .9639
             /estimated_length=unknown
misc_feature 9640. .11245
             /note="assembly_name:Contig34"
gap          11246. .11345
             /estimated_length=unknown
misc_feature 11346. .12748
             /note="assembly_name:Contig35"
gap          12749. .12848

```

Query Match 88.0%; Score 26.4; DB 14; Length 163399;

Best Local Similarity 96.4%; Pred. No. 8;

Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGCGTGTGTTAAAGACAA 28

Db 120846 CGGAATAAAGCGTGTGTTAAAGACAA 120819

RESULT 8

LOCUS

DEFINITION AL645608 186759 bp DNA linear PRI 18-MAY-2005
Human DNA sequence from clone RP11-5407 on chromosome 1 Contains
three novel genes, a novel gene (MGC45873), a novel gene
(DKFZ9564C186), a novel gene, a novel gene (DKFZP434H2010), a novel
gene (MGC13275), the gene for bHLH factor Hes4 (Hes4), a ribosomal
protein L39 (RPL39) pseudogene, a novel gene, the GIP2 gene for
interferon alpha-inducible protein (clone IFI-15K), the AGRN gene
for agrin, two novel genes and fifteen CpG islands, complete
sequence.

ACCESSION AL645608

VERSION AL645608.30 GI:48374144

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 186759)

REFERENCE

AUTHORS

TITLE Direct Submission

JOURNAL

Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk

On Jun 6, 2004 this sequence version replaced gi:45125204.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:

Em, ENBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information

on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human

chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr1>
RP11-5407 is from the library RP11-11.1 constructed by the group of
Pietér de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pBACE3.6

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: vega@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.

FEATURES

source

1. .186759

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="1"

/clone="RP11-5407"

/clone_lib="RP11-11.1"

954

/note="Clone right end: RP11-206L10"

join(<955..956,63125..63310,64994..65156,65986..66101,

66260..66338,66409..66908,67103..67227,67548..68109)

/locus_tag="RP11-5407.4-004"

join(<955..956,63125..63310,64994..65156,65986..66101,

66260..66338,66409..66908,67103..67227,67548..68109)

/locus_tag="RP11-5407.4-004"

/product="novel protein"

/note="match: CDNAS: Em:AK054643.1"

join(<955..956,63125..63310,64994..65156,65986..66101,

66260..66338,66409..66908,67103..67227,67548..68003)

/locus_tag="RP11-5407.4-004"

/standard_name="OTTHUMP00000038430"

/note="match: proteins: Q96NU1"

/codon_start=2

/product="novel protein"

/protein_id="CAI15563.1"

/db_xref="GI:55962171"

/db_xref="InterPro:IPR001660"

/db_xref="InterPro:IPR011310"

/db_xref="UniProt/TREMBL:O5SVA0"

/translation="EALLPRELGSNAPEDHYRLVLSALSEASTFEDPQRLYHLGLP

SHGEDPPHDPHPLPSHDLRVQEVAAALRGPSGLEAHLPSSTAGQRRKQGLAQH

REGAAPAAAPFSERELPQPPLLSPONAPHVALGPHLRPFLQVPSALCQTPGQFL

PQAQEMFANQQLRKQNLARLELPADLLRQKELASARQLLAPETALRPNDGAEEL

MTGARLVLNHAAPLLALPPQPGSPPTPSRDSARRAPKGGPGAPASRPSEKE

PLGFYVAVSYFHTGAVGLSMGDEAPAPEDVTWKTVDDVCSFVGLSGCGEYTRVP

REGIDGETLPLTEERHLTMGLKLPALKIAQVRRWGVRSQPHSVAESGWCVC

DSHQALSLQVARRLGRVFVYASFPVALPQPTPLRAPERELGTGEQPLSPPTATSPY

DGGHALAQTSKQENGTALLPGAPDPSQPLC"

join(<30099..30107,59622..59746,62890..62979,63125..63262,

64994..65156,65986..66101,66260..66338,66409..66908,

67103..67227,67548..67658,67758..68424)

/locus_tag="RP11-5407.4-001"

join(<30039..30107,59622..59746,62890..62979,63125..63262,

64994..65156,65986..66101,66260..66338,66409..66908,

67103..67227,67548..67658,67758..68424)

/locus_tag="RP11-5407.4-001"

/product="novel protein"

/note="match: ESTs: Em:BE258987.1 Em:CA389330.1

match: CDNAS: Em:BC024295.2"

join(<30099..30107,59622..59746,62890..62979,63125..63262,

64994..65156,65986..66101,66260..66338,66409..66908,

67103..67227,67548..67658,67758..68003)

/locus_tag="RP11-5407.4-001"

CDS

/standard_name="OTTHUMP0000038427"
/notes="match: proteins: Tr:Q8TB59"
/codon_start=2
/product="novel protein"
/protein_id="CAI15584.1"
/db_xref="GI:55962172"
/db_xref="InterPro:IPR001660"
/db_xref="InterPro:IPR011510"
/db_xref="UniProt/TREMBL:O5SV99"
/translation="SCSDGNLSTLTSSVSRHLYMPEHQSRCEQRGSLEIGLRPA
GDLLKRLGSDCFSEKRAESPOEALLPRELGPSPLEADHRYRLVSAUSE
ASTFEDQRIYHLGSLDRLVRQEVAAALRPGSGLEAHLFSSTAGRRKQGLAQH
REGFAAALPFPQPLPSQNAAPHVALGPHLRFPLGVSALCQTPGYGFL
PQAQENFAWQELLRKQNLARLEPADLLRQKLSARPOLLAPELTALRPNDGAEL
ORGLALVNHGAAPLALPQGPSPGPPTPSRDSARRAPKGGPGPASARSESKE
MTGNLWAQDSEDEPKDSGDGDPETAACVCGRPTFGQAPAGAGAEKGLFPQSTL
PLGFPYAFYFHGAVGSLGMBEAPEDVTKWTDVDFSGVGLSGCGGYITRVF
RQGIQDGTLLTEEHLTMGLKGPALKIRAVARRLGRVYVASFPVALPQQP
TLRAPRELGTGQPLSPPTTATSPYGGGHALAGTSPKQENGTLALLPGAPDPSQLP
"
32390. .32684
/locus_tag="RP11-5407.4-001"
/notes="Single clone region. Sequence generated from a
short insert library derived from a single pUC clone.
Restriction digest data confirm the assembly."
join(35285. .35323,35795. .38798)
/locus_tag="RP11-5407.1-001"
join(35285. .35323,35795. .38798)
/locus_tag="RP11-5407.1-001"
/product="novel transcript"
/notes="match: cDNAs: Em:AK056486.1"
join(38021. .38294,38653. .38821)
/locus_tag="RP11-5407.2-001"
join(38021. .38294,38653. .38821)
/locus_tag="RP11-5407.2-001"
/product="putative novel transcript"
/notes="match: ESTs: Em:AA63398.1"
complement(join(40720. .41570,41872. .41960,42675. .42765,
43185. .43542))
/locus_tag="RP11-5407.3-001"
complement(join(40720. .41570,41872. .41960,42675. .42765,
43185. .43542))
/locus_tag="RP11-5407.3-001"
/product="novel transcript"
/notes="match: ESTs: Em:BU619380.1 Em:BU632785.1
Em:BU738526.1 Em:CA447406.1"
complement(40720)
/locus_tag="RP11-5407.3-001"
complement(40741. .40746)
/locus_tag="RP11-5407.3-001"
complement(join(41220. .41336,41872. .42384))
/locus_tag="RP11-5407.3-002"
complement(join(41220. .41336,41872. .42384))
/locus_tag="RP11-5407.3-002"
/product="novel transcript"
/notes="match: ESTs: Em:BM758684.1 Em:BM759761.1
Em:BM765468.1"
complement(41423)
/locus_tag="RP11-5407.3-002"
complement(41440. .41445)
/locus_tag="RP11-5407.3-002"
join(<49000. .49039,49772. .49863,54005. .54186,54889. .54939,
59622. .>59643)
/locus_tag="RP11-5407.4-003"
/locus_tag="RP11-5407.4-003"
/product="novel protein"
/notes="match: ESTs: Em:AA427964.1"
join(<49000. .49039,49772. .49863,54005. .54186,54889. .54939,
59622. .>59643)
/locus_tag="RP11-5407.4-003"
/standard_name="OTTHUMP0000038429"

/codon_start=1
/product="novel protein"
/protein_id="CAI15565.1"
/db_xref="GI:55962173"
/db_xref="UniProt/TREMBL:O5SV95"
/translation="QRLGLAGSLRSEKSLKTLMSKGLQVHPPICDPCGCRISPV
NRGRADKRTVALPAARNLKGRTPSFASDGDSDGSGPTCGRRPGLKQDGPHIRIM
KRRVTHWDVNISFREASCSQGNLPT"
join(<49601. .49650,49772. .49863,54005. .54186,54889. .54939,
59622. .59746,62890. .>62972)
/locus_tag="RP11-5407.4-002"
join(<49601. .49650,49772. .49863,54005. .54186,54889. .54939,
59622. .59746,62890. .>62972)
/locus_tag="RP11-5407.4-002"
/product="novel protein"
/notes="match: ESTs: Em:CA396407.1"
join(<49601. .49650,49772. .49863,54005. .54186,54889. .54939,
59622. .59746,62890. .>62972)
/locus_tag="RP11-5407.4-002"
/standard_name="OTTHUMP0000038428"
/codon_start=2
/product="novel protein"
Query Match 88.0%; Score 26.4; DB 8; Length 186759;
Best Local Similarity 96.4%; Pred. No. 7.8;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGAATAAAGCGTGTGTTAAAGACAAA 28
|||||
DB 138363 CGGAATAAAGCGTGTGTTAAAGAGAAA 138390
|||||
RESULT 9
AB107102
LOCUS
DEFINITION
Homo sapiens chromosome 1 clone RP11-671C15 map lp36.3, **
SEQUENCING IN PROGRESS ***, 13 unordered pieces.
AB107102
VERSION
AB107102.1 GI:33636391
HTG; HTGS_PHASE1.
KEYWORDS
Homo sapiens
SOURCE
Homo sapiens
ORGANISM
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1
AUTHORS
Hosoya, N., Ogawa, S., Motokura, T., Hangaishi, A., Wang, L., Qiao, Y.,
Nannya, Y., Kogi, M. and Hirai, H.
TITLE
Molecular cytogenetic analyses of HIG, a novel human cell line
carrying t(1;3)(p36.3;q25.3) established from a patient with
chronic myelogenous leukemia in blasttic crisis
Int. J. Hematol. 78 (5), 432-438 (2003)
JOURNAL
PUBMED
14704036
REFERENCE
2 (bases 1 to 200491)
Hosoya, N.
AUTHORS
Direct Submission
TITLE
Submitted (28-MAR-2003) Noriko Hosoya, University of Tokyo.
JOURNAL
Department of Hematology and Oncology, Hongo 7-3-1, Bunkyo-ku,
Tokyo 113-8655, Japan (E-mail:hosoya-tky@umin.ac.jp,
Tel:81-3-3815-5411(ex.35602), Fax:81-3-5804-6261)
COMMENT
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 35799: contig of 35799 bp in length
* 35800 35899: gap of 100 bp
* 35900 39521: contig of 3622 bp in length
* 39522 39621: gap of 100 bp
* 39622 53971: contig of 14350 bp in length

* 53972 54071: gap of 100 bp
 * 54072 89894: contig of 35823 bp in length
 * 89895 89994: gap of 100 bp
 * 89995 109993: contig of 19999 bp in length
 * 109994 110093: gap of 100 bp
 * 110094 124472: contig of 14379 bp in length
 * 124473 124572: gap of 100 bp
 * 124573 128857: contig of 4285 bp in length
 * 128858 128957: gap of 100 bp
 * 128958 139050: contig of 10093 bp in length
 * 139051 139150: gap of 100 bp
 * 139151 145733: contig of 6583 bp in length
 * 145734 145833: gap of 100 bp
 * 145834 156433: contig of 10600 bp in length
 * 156434 156533: gap of 100 bp
 * 156534 177511: contig of 20978 bp in length
 * 177512 177611: gap of 100 bp
 * 177612 186275: contig of 8664 bp in length
 * 186276 186375: gap of 100 bp
 * 186376 200491: contig of 14116 bp in length.

FEATURES

source
 1..200491
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1p36.3"
 /clone="RP11-671C15"

ORIGIN

Query Match 88.0%; Score 26.4; DB 14; Length 200491;
 Best Local Similarity 96.4%; Pred. No. 7.7;
 Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGCGCTGTTGTTAAAGACAAA 28
 |||||
 Db 150670 CGGAATAAAGCGCTGTTGTTAAAGACAAA 150697

RESULT 10

BX571684
 LOCUS Homo sapiens chromosome 1 clone CTD-3032124, 3 unordered pieces.
 DEFINITION BX571684
 ACCESSION BX571684.2 GI:32812682
 VERSION HTGS_PHASE1; HTGS_CANCELLED.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo
 1 (bases 1 to 327883)
 Hall, R.

REFERENCE

AUTHORS Direct Submission
 TITLE Submitted (14-JUL-2003) Wellcome Trust Sanger Institute, Hinxton,
 JOURNAL Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

COMMENT

On Jul 15, 2003 this sequence version replaced gi:32567439.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk
 ----- Project Information
 Center project name: BX3032124
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Chemistry: Dye-terminator; 100% of reads
 Consensus quality: 184256 bases at least Q40
 Consensus quality: 184906 bases at least Q30
 Consensus quality: 185360 bases at least Q20
 Insert size: 327683; sum-of-contigs
 Insert size: 195389; 4.5% error; agarose-fp

Quality coverage: 4.02x in Q20 bases; sum-of-contigs Quality coverage: 6.85x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

* 1 194085: contig of 194085 bp in length
 * 194086 194185: gap of 100 bp
 * 194186 204837: contig of 10652 bp in length
 * 204838 204937: gap of 100 bp
 * 204938 327883: contig of 122946 bp in length.

FEATURES

source
 1..327883
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="CTD-3032124"
 /clone_lib="CIT-HSP-D2"
 1..194085
 /note="assembly_fragment:02147"
 fragment_chain:1
 194186..204837
 /note="assembly_fragment:01387"
 fragment_chain:1
 204938..327883
 /note="assembly_fragment:02101"

ORIGIN

Query Match 88.0%; Score 26.4; DB 14; Length 327883;
 Best Local Similarity 96.4%; Pred. No. 6.9;
 Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGCGCTGTTGTTAAAGACAAA 28
 |||||
 Db 139484 CGGAATAAAGCGCTGTTGTTAAAGACAAA 139511

RESULT 11

LOCUS Homo sapiens (human)
 DEFINITION Sequence 4170 from Patent WO2004097052.
 ACCESSION CQ922970
 VERSION CQ922970.1 GI:56212911
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo
 1

REFERENCE

AUTHORS Strachan, A., Immerman, P. and Dörner, A. J.
 TITLE Methods for prognosis and treatment of solid tumors
 JOURNAL Patent: WO 2004097052-A 4170 11-NOV-2004;
 Wyeth (US); Burczynski, Michael E. (US)

FEATURES

source
 1..509
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 100
 /note="n includes a, c, g, or t, or contains no nucleotide"
 110
 /note="n includes a, c, g, or t, or contains no nucleotide"
 163
 misc_feature

misc_feature /note="n includes a, c, g, or t, or contains no nucleotide" 185
misc_feature /note="n includes a, c, g, or t, or contains no nucleotide" 257..260
misc_feature /note="n includes a, c, g, or t, or contains no nucleotide" 300
misc_feature /note="n includes a, c, g, or t, or contains no nucleotide" 307
misc_feature /note="n includes a, c, g, or t, or contains no nucleotide" 348..367
misc_feature /note="n includes a, c, g, or t, or contains no nucleotide" 422..483
misc_feature /note="n includes a, c, g, or t, or contains no nucleotide" 307

ORIGIN

Query Match 81.3%; Score 24.4; DB 6; Length 509;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTGTTAAGACAAAAA 30
|||||
Db 236 CGGAATAAAGGCTGTGTTAANNNGAAAA 265

RESULT 12

AP000876/c
LOCUS AP000876 161151 bp DNA linear HTG 30-MAY-2000
DEFINITION Homo sapiens chromosome 18 clone RP11-771B1 map 18p11.3, WORKING DRAFT SEQUENCE, 17 unordered pieces.

ACCESSION AP000876
VERSION AP000876.2 GI-8119024
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 161151)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

Homo sapiens 161,151 genomic DNA of 18p11.3
Published Only in Database (1999)
2 (bases 1 to 161151)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

Direct Submission
Submitted (13-DEC-1999) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan (E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-42-778-9923, Fax: 81-42-778-9924)

On May 31, 2000 this sequence version replaced gi:6997705.
----- Genome Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft18
Center clone name: RP11-771B1
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 147768 bases at least Q40

Consensus quality: 154700 bases at least Q30
Consensus quality: 157838 bases at least Q20
Insert size: 159551; sum-of-contigs
Quality coverage: 4.74x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 24287 contig of 24287 bp in length
24388 4446 contig of 20059 bp in length
59447 59040 contig of 14494 bp in length
59441 74881 contig of 15741 bp in length
74982 89942 contig of 14961 bp in length
90043 103712 contig of 13670 bp in length
103813 114497 contig of 10685 bp in length
114598 125844 contig of 11247 bp in length
125945 133701 contig of 7757 bp in length
133802 140062 contig of 6261 bp in length
140163 145470 contig of 5308 bp in length
145571 150374 contig of 4804 bp in length
150475 154892 contig of 4418 bp in length
154893 155990 contig of 998 bp in length
156091 157891 contig of 1801 bp in length
157892 159471 contig of 1480 bp in length
159472 159571 contig of 100 bp in length
159572 161151 contig of 1580 bp in length

Sequence updated (26-May-2000).

* NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 24287: contig of 24287 bp in length
24288 24387: gap of 100 bp
24388 4446: contig of 20059 bp in length
44447 44546: gap of 100 bp
44547 59040: contig of 14494 bp in length
59041 59140: gap of 100 bp
59141 74881: contig of 15741 bp in length
74882 74981: gap of 100 bp
74982 89942: contig of 14961 bp in length
89943 90042: gap of 100 bp
90043 103712: contig of 13670 bp in length
103713 103812: gap of 100 bp
103813 114497: contig of 10685 bp in length
114498 114597: gap of 100 bp
114598 125844: contig of 11247 bp in length
125945 125944: gap of 100 bp
125945 133701: contig of 7757 bp in length
133702 133801: gap of 100 bp
133802 140062: contig of 6261 bp in length
140063 140162: gap of 100 bp
140163 145470: contig of 5308 bp in length
145471 145570: gap of 100 bp
145571 150374: contig of 4804 bp in length
150375 150474: gap of 100 bp
150475 154892: contig of 4418 bp in length
154893 154990: contig of 998 bp in length
154993 155990: contig of 998 bp in length
155991 156090: gap of 100 bp
156091 157891: contig of 1801 bp in length
157892 157991: gap of 100 bp
157992 159471: contig of 1480 bp in length
159472 159571: gap of 100 bp
159572 161151: contig of 1580 bp in length.

Location/Qualifiers

1..161151

FEATURES
source


```
repeat_region /rpt_family="MLT1I"
15540..15705
/rpt_family="MER5B"
15836..16582
/rpt_family="L1PA6"
complement(16583..17089)
/rpt_family="MER9"
17090..22509
/rpt_family="L1PA6"
22774..23406
/rpt_family="L1MEC"
complement(23983..24178)
/rpt_family="MLT1J"
complement(25899..26406)
/rpt_family="MLT1D"
26790..26810
/rpt_family="(CCTG)n"
complement(30033..31005)
/rpt_family="L2"
31546..31959
/rpt_family="MLT1D"
31960..32236
/rpt_family="MER44D"
complement(32237..32591)
/rpt_family="MER47A"
32592..32991
/rpt_family="MER44D"
complement(34134..34349)
/rpt_family="MIR"
34424..34847
/rpt_family="L1MC4"
complement(34871..34969)
/rpt_family="AluJo/FRAM"
35021..35144
/rpt_family="(TATATG)n"
35258..35358
/rpt_family="AT_rich"
35821..35658
/rpt_family="AT_rich"
complement(35783..35897)
/rpt_family="MIR"
complement(36089..36316)
/rpt_family="MIR"
36787..37087
/rpt_family="AluY"
37383..37556
/rpt_family="MER102"
37578..37607
/rpt_family="(CA)n"
39358..39383
/rpt_family="(T)n"
39936..40082
/rpt_family="MIR"
40495..40637
/rpt_family="MER5B"
41783..42073
/rpt_family="AluSg"
42856..42965
/rpt_family="MIR"
45120..45328
/rpt_family="L1MB8"
46246..46398
/rpt_family="MER5B"
complement(46520..46819)
/rpt_family="AluSg"
complement(47951..48228)
/rpt_family="MER33"
50241..50533
/rpt_family="AluJo"
51158..51215
/rpt_family="(TA)n"
51624..52514
/rpt_family="Tigger3b"
```

```
repeat_region 52515..52741
/rpt_family="MER30"
52742..52934
/rpt_family="Tigger3b"
53342..53375
/rpt_family="AT_rich"
53450..53754
/rpt_family="AluSx"
54957..55256
/rpt_family="AluYb8"
55888..55921
/rpt_family="AT_rich"
56168..56581
/rpt_family="MSTA"
56833..56920
/rpt_family="L1MA9"
57982..58106
/rpt_family="MER103"
58703..59203
/rpt_family="L1MC5"
complement(59212..59426)
/rpt_family="MER20"
59443..59650
/rpt_family="L1MC5"
59651..59937
/rpt_family="AluSp"
59938..60333
/rpt_family="L1MC5"
61866..61886
/rpt_family="AT_rich"
complement(62370..62523)
/rpt_family="AluJo"
62524..62557
/rpt_family="(TTTG)n"
```

Query Match 78.0%; Score 23.4; DB 8; Length 211945;
Best Local Similarity 96.0%; Pred. No. 94;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 ATAAAGCGTGTGTTAAAGACAAAAA 30

DB 129603 ATAAAGCGTGTGTTAAAGATAAAAA 129627

RESULT 14

BD275867/c 301 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS OF LUNG CANCER.
ACCESSION BD275867
VERSION BD275867.1 GI:33085635
KEYWORDS JP 2002543769-A/261.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 301)
AUTHORS Wang,T. and Van,R.
TITLE COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS OF LUNG CANCER
JOURNAL Patent: JP 2002543769-A 261 24-DEC-2002;
Corixa Corporation et al
COMMENT OS Homo sapien
PN JP 2002543769-A/261
PD 24-DEC-2002
PF 03-APR-2000 JP 2000611554
PR 22-FEB-2000 US 09/510376,10-JAN-2000 US 09/480884, PR
02-APR-1999 US 09/285479,30-DEC-1999 US 09/476496, PR
17-DEC-1999 US 09/466396
PI tongtong wang,rikn van
CC
FH Key Location/Qualifiers
FT misc_feature (1)..(301)
FT /note='n = A,T,C or G'.
FEATURES Location/Qualifiers
source 1..301

/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 76.7%; Score 23; DB 6; Length 301;
Best Local Similarity 95.8%; Pred. No. 5.6e+02;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTTAAAGA 24
|||
Db 25 CGNAATAAAGGCTGTTGTTAAAGA 2

RESULT 15

AR220652/c 301 bp DNA linear PAT 26-SEP-2002

LOCUS AR220652
DEFINITION Sequence 304 from patent US 6426072.
ACCESSION AR220652
VERSION AR220652.1 GI:23327433

KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 301)
AUTHORS Wang, T., Fan, L., Kalos, M.D., Bangur, C.S., Hosken, N.A., Fanger, G.R.,
Li, S.X., Wang, A., Skeiky, Y.A.W., Henderson, R.A. and McNeill, P.D.
TITLE Compositions and methods for the therapy and diagnosis of lung
cancer

JOURNAL Patent: US 6426072-A 304 30-JUL-2002;
Corixa Corporation; Seattle, WA

FEATURES
source 1..301
Location/Qualifiers
1..301
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 76.7%; Score 23; DB 6; Length 301;
Best Local Similarity 95.8%; Pred. No. 5.6e+02;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTTAAAGA 24
|||
Db 25 CGNAATAAAGGCTGTTGTTAAAGA 2

Search completed: December 7, 2005, 04:22:16
Job time : 270.316 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:31:40 ; Search time 74.7009 Seconds
(without alignments)
3320.997 Million cell updates/sec

Title: US-09-980-046B-9
Perfect score: 30
Sequence: 1 cggaaataaaggctgttgaagacaaaaa 30

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	28.4	94.7	547	3	US-09-997-003-18
2	28.4	94.7	642	5	US-10-044-090-536
3	28.4	94.7	670	6	US-10-131-410-2
4	28.4	94.7	678	8	US-10-723-860-695
5	28.4	94.7	678	9	US-10-756-149-705
6	28.4	94.7	1323	6	US-10-264-049-307
7	28.4	94.7	1389	8	US-10-723-860-5322
8	23	76.7	301	3	US-09-733-703-304
9	23	76.7	301	3	US-09-850-716A-304
10	23	76.7	301	3	US-09-897-778-304
11	23	76.7	301	5	US-10-007-700-304
12	23	76.7	301	6	US-10-117-982-304
13	23	76.7	301	6	US-10-313-986-304
14	23	76.7	301	8	US-10-775-972-304
15	23	76.7	301	9	US-10-922-124-304
16	21.2	70.7	905	5	US-10-198-846-4040
17	21	70.0	464	3	US-09-764-887-132
18	21	70.0	464	5	US-10-073-961-122
19	21	70.0	486	3	US-09-764-887-586
20	21	70.0	486	3	US-09-764-887-587
21	21	70.0	486	5	US-10-073-961-586
22	21	70.0	486	5	US-10-073-961-587
23	21	70.0	634	3	US-09-954-531-144

Sequence 362, App
Sequence 571, App
Sequence 233, App
Sequence 190, App
Sequence 140, App
Sequence 1293, App
Sequence 105, App
Sequence 231, App
Sequence 214, App
Sequence 4, Appli
Sequence 195, App
Sequence 95, Appl
Sequence 1211, App
Sequence 1429, App
Sequence 1638, App
Sequence 67, Appl
Sequence 16152, A
Sequence 48862, A
Sequence 48904, A
Sequence 48920, A
Sequence 732924,
Sequence 732979,

24 21 70.0 634 3 US-09-954-531-362
25 21 70.0 634 3 US-09-954-531-571
26 21 70.0 634 6 US-10-177-293-233
27 21 70.0 634 6 US-10-101-510-190
28 21 70.0 634 6 US-10-159-563-140
29 21 70.0 634 7 US-10-240-425-1293
30 21 70.0 634 7 US-10-188-832-105
31 21 70.0 634 7 US-10-755-889-231
32 21 70.0 634 8 US-10-775-169-214
33 21 70.0 634 8 US-10-802-432-4
34 21 70.0 634 8 US-10-370-715B-195
35 21 70.0 634 8 US-10-487-422-95
36 21 70.0 634 9 US-10-843-641A-1211
37 21 70.0 634 9 US-10-843-641A-1429
38 21 70.0 634 9 US-10-843-641A-1638
39 21 70.0 634 9 US-10-631-467-67
40 20.6 68.7 148 8 US-10-674-124A-16152
41 20.6 68.7 201 8 US-10-719-993-48862
42 20.6 68.7 201 8 US-10-719-993-48904
43 20.6 68.7 201 8 US-10-719-993-48920
44 20.6 68.7 538 4 US-09-925-065A-732924
45 20.6 68.7 575 4 US-09-925-065A-732979

ALIGNMENTS

RESULT 1
US-09-997-003-18
; Sequence 18, Application US/09997003
; Publication No. US20030203361A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA003P1
; CURRENT APPLICATION NUMBER: US/09/997,003
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: PCT/US00/22157
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/148,680
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 547
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-997-003-18

Query Match 94.7%; Score 28.4; DB 3; Length 547;
Best Local Similarity 96.7%; Pred. No. 0.27; Mismatches 0; Indels 1; Gaps 0;
Matches 29; Conservative 0

QY 1 CGGAATAAAGGCTGTTGTAAGACAAAAA 30
|||||
Db 488 CGGAATAAAGGCTGTTGTAAGACAAAAA 517

RESULT 2
US-10-044-090-536
; Sequence 536, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 536

```
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1307829CB1
US-10-044-090-536

Query Match          94.7%; Score 28.4; DB 5; Length 642;
Best Local Similarity 96.7%; Pred. No. 0.28;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTTGTTAAAGACAAAAA 30
   |||||||||||||||||||||||||||||||
Db 606 CGGAATAAAGGCTGTTGTTAAAGACAAAAA 635

RESULT 3
; Sequence 2, Application US/10131410
; Publication No. US20030235915A1
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST
; FILE REFERENCE: SCH-1763
; CURRENT APPLICATION NUMBER: US/10/131,410
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 09/646,673
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: PCT/DE99/00908
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 670
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-410-2

Query Match          94.7%; Score 28.4; DB 6; Length 670;
Best Local Similarity 96.7%; Pred. No. 0.28;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTTGTTAAAGACAAAAA 30
   |||||||||||||||||||||||||||||||
Db 637 CGGAATAAAGGCTGTTGTTAAAGACAAAAA 666

RESULT 4
; Sequence 695, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 695
; LENGTH: 678
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-695

Query Match          94.7%; Score 28.4; DB 8; Length 678;
Best Local Similarity 96.7%; Pred. No. 0.28;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTTGTTAAAGACAAAAA 30
   |||||||||||||||||||||||||||||||
Db 620 CGGAATAAAGGCTGTTGTTAAAGACAAAAA 649

RESULT 5
; Sequence 705, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 705
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-756-149-705

Query Match          94.7%; Score 28.4; DB 9; Length 678;
Best Local Similarity 96.7%; Pred. No. 0.28;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTTGTTAAAGACAAAAA 30
   |||||||||||||||||||||||||||||||
Db 620 CGGAATAAAGGCTGTTGTTAAAGACAAAAA 649

RESULT 6
; Sequence 307, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL13P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 307
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-264-049-307

Query Match          94.7%; Score 28.4; DB 6; Length 1323;
Best Local Similarity 96.7%; Pred. No. 0.32;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTTGTTAAAGACAAAAA 30
   |||||||||||||||||||||||||||||||
Db 659 CGGAATAAAGGCTGTTGTTAAAGACAAAAA 688
```


RESULT 7
US-10-723-860-5322
; Sequence 5322, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Azila, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5322
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-5322

Query Match 94.7%; Score 28.4; DB 8; Length 1389;
Best Local Similarity 96.7%; Pred. No. 0.32;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGACAAAA 30
|||
Db 554 CGGAATAAAGGCTGTTGTAAGAGAAAA 583

RESULT 8
US-09-735-705-304/c
; Sequence 304, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(301)
; OTHER INFORMATION: n = A,T,C or G
US-09-735-705-304

Query Match 76.7%; Score 23; DB 3; Length 301;
Best Local Similarity 95.8%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24
|||
Db 25 CGGAATAAAGGCTGTTGTAAGA 2

RESULT 9
US-09-850-716A-304/c
; Sequence 304, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(301)
; OTHER INFORMATION: n = A,T,C or G
US-09-850-716A-304

Query Match 76.7%; Score 23; DB 3; Length 301;
Best Local Similarity 95.8%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24
|||
Db 25 CGGAATAAAGGCTGTTGTAAGA 2

RESULT 10
US-09-897-778-304/c
; Sequence 304, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Barrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 23, 104, 192
; OTHER INFORMATION: n = A,T,C or G
US-09-897-778-304

Query Match 76.7%; Score 23; DB 3; Length 301;
Best Local Similarity 95.8%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24
|||
Db 25 CGGAATAAAGGCTGTTGTAAGA 2

```
Db      25  CGNAATAAAGGCTGTTGTAAGA 2

RESULT 11
US-10-007-700-304/c
; Sequence 304, Application US/10007700
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, AiJun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 23, 104, 192
; OTHER INFORMATION: n = A,T,C or G
US-10-007-700-304

Query Match      76.7%; Score 23; DB 5; Length 301;
Best Local Similarity 95.8%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  CGGAATAAAGGCTGTTGTAAGA 24
Db      25  CGNAATAAAGGCTGTTGTAAGA 2

RESULT 12
US-10-117-982-304/c
; Sequence 304, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 23, 104, 192
; OTHER INFORMATION: n = A,T,C or G
US-10-117-982-304

Query Match      76.7%; Score 23; DB 5; Length 301;
Best Local Similarity 95.8%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  CGGAATAAAGGCTGTTGTAAGA 24
Db      25  CGNAATAAAGGCTGTTGTAAGA 2

RESULT 13
US-10-313-986-304/c
; Sequence 304, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 23, 104, 192
; OTHER INFORMATION: n = A,T,C or G
US-10-313-986-304

Query Match      76.7%; Score 23; DB 6; Length 301;
Best Local Similarity 95.8%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  CGGAATAAAGGCTGTTGTAAGA 24
Db      25  CGNAATAAAGGCTGTTGTAAGA 2

RESULT 14
US-10-775-972-304/c
; Sequence 304, Application US/10775972
; Publication No. US20040235072A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C21
; CURRENT APPLICATION NUMBER: US/10/775,972
; CURRENT FILING DATE: 2004-02-10
; NUMBER OF SEQ ID NOS: 563
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 23, 104, 192
; OTHER INFORMATION: n = A,T,C or G
US-10-775-972-304/c

Query Match      76.7%; Score 23; DB 6; Length 301;
Best Local Similarity 95.8%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  CGGAATAAAGGCTGTTGTAAGA 24
Db      25  CGNAATAAAGGCTGTTGTAAGA 2
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 23, 104, 192
; OTHER INFORMATION: n = A,T,C or G
US-10-775-972-304

Query Match          76.7%; Score 23; DB 8; Length 301;
Best Local Similarity 95.8%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGCGCTGTTGTAAGA 24
Db 25 CGNAATAAAGCGCTGTTGTAAGA 2

RESULT 15
US-10-922-124-304/c
; Sequence 304, Application US/10922124
; Publication No. US20050142620A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Zehentner-Wilkinson, Barbara K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C22
; CURRENT APPLICATION NUMBER: US/10/922,124
; CURRENT FILING DATE: 2004-08-18
; PRIOR APPLICATION NUMBER: 10/775,972
; PRIOR FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: 10/623,155
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/313,986
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 10/117,982
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 10/007,700
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 09/897,778
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 09/850,716
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 09/735,705
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 09/685,696
; PRIOR FILING DATE: 2000-10-09
; PRIOR APPLICATION NUMBER: 09/662,786
; PRIOR FILING DATE: 2000-09-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 563
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 23, 104, 192
; OTHER INFORMATION: n = A,T,C or G
US-10-922-124-304

Query Match          76.7%; Score 23; DB 9; Length 301;
Best Local Similarity 95.8%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGCGCTGTTGTAAGA 24
Db 25 CGNAATAAAGCGCTGTTGTAAGA 2

Search completed: December 6, 2005, 22:57:12
Job time : 76.7009 secs
```

This Page Blank (Lap 1)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 970.211 Seconds
(without alignments)
4099.008 Million cell updates/sec

Title: US-09-980-046B-10
Perfect score: 85
Sequence: 1 cggfgaagtgcattcttcgc.....tgccccactaagttagaaaaa 85

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_hc:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_est7:
9: gb_gsa1:
10: gb_gsa2:
11: gb_gsa3:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	83.4	98.1	155	AA235328	AA235328 zt30c08.s
C 2	83.4	98.1	186	AA292318	AA292318 zt30c08.r
C 3	83.4	98.1	243	AA631632	AA631632 np59605.s
C 4	83.4	98.1	248	BF590563	BF590563 7h38e05.x
C 5	83.4	98.1	268	AA514303	AA514303 nf60c10.s
C 6	83.4	98.1	270	AI334968	AI334968 tb21b05.x
C 7	83.4	98.1	272	AI307267	AI307267 tb24d10.x
C 8	83.4	98.1	272	BF437551	BF437551 7p74a08.x
C 9	83.4	98.1	273	AI400926	AI400926 th25e10.x
C 10	83.4	98.1	273	AI669789	AI669789 tu31c11.x
C 11	83.4	98.1	300	AI718891	AI718891 as76e03.x
C 12	83.4	98.1	303	BE646272	BE646272 7e84a01.x
C 13	83.4	98.1	304	AI417902	AI417902 t955d11.x
C 14	83.4	98.1	306	AI760920	AI760920 w170d11.x
C 15	83.4	98.1	309	AA148966	AA148966 zn99g11.s
C 16	83.4	98.1	315	AI470973	AI470973 tf89a08.x
C 17	83.4	98.1	321	AI468348	AI468348 tg58a07.x
C 18	83.4	98.1	325	CK430031	CK430031 oJ44c02.y
C 19	83.4	98.1	327	AI867674	AI867674 wd78f04.x
C 20	83.4	98.1	329	BF063247	BF063247 7h87b01.x
C 21	83.4	98.1	336	AI708880	AI708880 as86e10.x
C 22	83.4	98.1	338	AI379991	AI379991 tg18e08.x

C 23	83.4	98.1	354	1	AW768412	AW768412 hk74e04.x
C 24	83.4	98.1	355	1	AW627370	AW627370 hh83f07.x
C 25	83.4	98.1	362	1	AA974311	AA974311 oq10e12.s
C 26	83.4	98.1	365	2	BF059613	BF059613 7k63e06.x
C 27	83.4	98.1	365	8	RI2367	RI2367 yf28a06.s
C 28	83.4	98.1	379	5	BU963341	BU963341 AGENCOURT
C 29	83.4	98.1	384	1	AA938338	AA938338 oo98a03.s
C 30	83.4	98.1	387	6	CA436188	CA436188 UI-H-DT0-
C 31	83.4	98.1	391	1	AI673220	AI673220 we72e09.x
C 32	83.4	98.1	395	1	AI659337	AI659337 tu29d03.x
C 33	83.4	98.1	397	1	AI368956	AI368956 qw15h02.x
C 34	83.4	98.1	402	3	BQ011944	BQ011944 UI-1-BC1P
C 35	83.4	98.1	405	1	AW452170	AW452170 UI-H-B13-
C 36	83.4	98.1	417	1	AI819871	AI819871 wj05a04.x
C 37	83.4	98.1	420	1	AI028759	AI028759 ov98f12.x
C 38	83.4	98.1	423	1	AI673198	AI673198 we72c09.x
C 39	83.4	98.1	429	6	CF529461	CF529461 UI-1-BC1P
C 40	83.4	98.1	431	2	BF003132	BF003132 7g55f07.x
C 41	83.4	98.1	443	1	AI870930	AI870930 w169e11.x
C 42	83.4	98.1	443	3	BQ011066	BQ011066 UI-1-BC1P
C 43	83.4	98.1	445	1	AI278959	AI278959 qo51h12.x
C 44	83.4	98.1	448	1	AA100479	AA100479 z181e09.s
C 45	83.4	98.1	448	8	W74705	W74705 zd57e08.s

ALIGNMENTS

RESULT 1
AA235328/c

LOCUS
DEFINITION
IMAGE:723854 3' similar to gb:L11045_rna1 PROTEASOME COMPONENT C13 (HUMAN); mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AA235328 155 bp mRNA linear EST 08-AUG-1997
zt30c08.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
IMAGE:723854 3' similar to gb:L11045_rna1 PROTEASOME COMPONENT C13 (HUMAN); mRNA sequence.
ACCESSION AA235328.1 GI:1859965
VERSION AA235328.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 155)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chissole, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, P., Thierry-Mieg, J.,
Trevasakis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.
and Maria, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
8889549
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 254 Std Error: 0.00
Seq primer: -41m13 fwd. ET from Amerham.
Location/Qualifiers
1. .155
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:723854"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares ovary tumor NbHOT"

/note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGGTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 98.1%; Score 83.4; DB 1; Length 155;
Best Local Similarity 98.8%; Pred. No. 3.5e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAACAAGCATTTCCCGGAGGAACTTT 60
Db 99 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAACAAGCATTTCCCGGAGGAACTTT 40

Qy 61 CTGGTGCCCCCACTAAGTAGAAAA 85
Db 39 CTGGTGCCCCCACTAAGTAGAATAA 15

RESULT 2
AA292318
LOCUS
DEFINITION
AA292318.1 Homo sapiens ovary tumor N8HOT Homo sapiens cDNA clone IMAGE:723854 5' similar to gb:L11045_rnal PROTEASOME COMPONENT C13 (HUMAN); mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Putative full length read
The vector to vector length is 187
Insert Length: 254 Std Error: 0.00
Seq primer: -28ml3 rev2 ET from Amersham.

FEATURES
Location/Qualifiers
1..186
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:723854"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares ovary tumor N8HOT"
/note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st

strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGGTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 98.1%; Score 83.4; DB 1; Length 186;
Best Local Similarity 98.8%; Pred. No. 3.6e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAACAAGCATTTCCCGGAGGAACTTT 60
Db 89 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAACAAGCATTTCCCGGAGGAACTTT 148

Qy 61 CTGGTGCCCCCACTAAGTAGAAAA 85
Db 149 CTGGTGCCCCCACTAAGTAGAATAA 173

RESULT 3
AA631632/c
LOCUS
DEFINITION
AA631632.1 Homo sapiens Bz2 Homo sapiens cDNA clone IMAGE:1130624 3' similar to gb:L11045_rnal PROTEASOME COMPONENT C13 (HUMAN); mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: csap@ncic.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.lnl.gov/bbrp/image/image.html
Insert Length: 651 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham.

FEATURES
Location/Qualifiers
1..243
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1130624"
/sex="female, pooled"
/tissue_type="breast"
/lab_host="DH10B"
/clone_lib="NCI-CCGAP Bz2"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled bulk breast tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. This library is the normalized version of NCI CCGAP Bz1.1. Library was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 98.1%; Score 83.4; DB 1; Length 243;
 Best Local Similarity 98.8%; Pred. No. 3.9e-18;
 Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAAACAGCATTTCCCGCCAGGGAAGTTT 60
 DB 103 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAAACAGCATTTCCCGCCAGGGAAGTTT 44

QY 61 CTGGTGCCCACTAAGTAGAATAA 85

DB 43 CTGGTGCCCACTAAGTAGAATAA 19

RESULT 4
 BF590563/c
 LOCUS
 DEFINITION 7h38e05.x1 NCI_CGAP_Col6 Homo sapiens cDNA clone IMAGE:3318272 3', mRNA sequence.

ACCESSION BF590563
 VERSION BF590563.1 GI:11682887

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE 1 (bases 1 to 248)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck,
 M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL, send email to:
 info@image.llnl.gov
 Seq primer: -40UP from Gibco.
 Location/Qualifiers

FEATURES

source

1..248
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3318272"
 /tissue_type="colon tumor, RER+"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Col6"
 /note="Organ: colon; Vector: p7T3D-Pac (Pharmacia) with a
 modified polylinker; Site 1: Not 1; Site 2: Eco RI;
 Plasmid DNA from the normalized library NCI_CGAP Col0 was
 prepared, and as circles were made in vitro. Following HAP
 purification, this DNA was used as tracer in a subtractive
 hybridization reaction. The driver was PCR-amplified cDNAs
 from a pool of 5,000 clones made from the same library
 (cloneIds 1057416-1061255, and 1144584-1145351).
 Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 98.1%; Score 83.4; DB 2; Length 248;
 Best Local Similarity 98.8%; Pred. No. 3.9e-18;
 Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAAACAGCATTTCCCGCCAGGGAAGTTT 60
 DB 104 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAAACAGCATTTCCCGCCAGGGAAGTTT 45

QY 61 CTGGTGCCCACTAAGTAGAATAA 85

DB 44 CTGGTGCCCACTAAGTAGAATAA 20

RESULT 5

RA514303/c

LOCUS

DEFINITION

AA514303

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck,
 M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 354 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham.
 Location/Qualifiers

1..268

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:924306"

/sex="pooled"

/tissue_type="colon"

/lab_host="DH10B"

/clone_lib="NCI CGAP Co3"

/note="Vector: p7T3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA
 was prepared from 12 pooled bulk tumor samples and primed
 with a Not I - oligo(dT) primer. Double-stranded cDNA was
 ligated to Eco RI adaptors (Pharmacia), digested with Not
 I and cloned into the Not I and Eco RI sites of the
 modified p7T3 vector. Library went through one round of
 normalization."

ORIGIN

Query Match 98.1%; Score 83.4; DB 1; Length 268;
 Best Local Similarity 98.8%; Pred. No. 3.9e-18;
 Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAAACAGCATTTCCCGCCAGGGAAGTTT 60
 DB 100 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAAACAGCATTTCCCGCCAGGGAAGTTT 41

QY 61 CTGGTGCCCACTAAGTAGAATAA 85

DB 40 CTGGTGCCCACTAAGTAGAATAA 16

RESULT 6

AI334968/c

LOCUS

DEFINITION

AI334968

tb2ib05.x1 NCI CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2054961 3',
 similar to gb:L11045_rnal PROTEASOME COMPONENT C13 (HUMAN); mRNA
 sequence.

```

ACCESSION      AI334968
VERSION        AI334968.1  GI:4071895
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Homnidae; Homo.

REFERENCE      1 (bases 1 to 270)
AUTHORS       NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE         National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL       Unpublished (1997)
COMMENT       Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-remail.nih.gov
               Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
               Emmert-Buck, M.D., Ph.D.
               cDNA Library Preparation: M. Bento Soares, Ph.D.
               DNA Sequencing by: Washington University Genome Sequencing Center
               Clone distribution: NCI-CGAP clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               www-bio.llnl.gov/bbrp/image/image.html
               Insert Length: 355 Std Error: 0.00
               Seq primer: -40UP from Gibco.
               Location/Qualifiers
               1..270
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:2054961"
                 /tissue_type="2 pooled tumors (clear cell type)"
                 /lab_host="DH10B"
                 /clone_lib="NCI CGAP Kid12"
                 /note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
                 a modified polylinker; Site 1: Not I; Site 2: Eco RI;
                 Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
                 prepared, and ss circles were made in vitro. Following HAP
                 purification, this DNA was used as tracer in a subtractive
                 hybridization reaction. The driver was PCR-amplified cDNAs
                 from a pool of 5,000 clones made from the same library
                 (cloneIDs 1323912-1325831, 1471368-1472903 and
                 1492104-1493255). Subtraction by Bento Soares and M.
                 Fatima Bonaldo."

FEATURES       source
               1..270
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:2054961"
                 /tissue_type="2 pooled tumors (clear cell type)"
                 /lab_host="DH10B"
                 /clone_lib="NCI CGAP Kid12"
                 /note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
                 a modified polylinker; Site 1: Not I; Site 2: Eco RI;
                 Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
                 prepared, and ss circles were made in vitro. Following HAP
                 purification, this DNA was used as tracer in a subtractive
                 hybridization reaction. The driver was PCR-amplified cDNAs
                 from a pool of 5,000 clones made from the same library
                 (cloneIDs 1323912-1325831, 1471368-1472903 and
                 1492104-1493255). Subtraction by Bento Soares and M.
                 Fatima Bonaldo."

ORIGIN
               Query Match      98.1%; Score 83.4; DB 1; Length 270;
               Best Local Similarity 98.8%; Pred. No. 3.9e-18;
               Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTCATCTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 110 CGGTGAAGTCATCTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 51

Qy 61 CTGGTGCCCACTAAGTAGAAAAA 85
Db 50 CTGGTGCCCACTAAGTAGAATAA 26

RESULT 7
AI307267/c
LOCUS         t234d10.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2055283 3'
DEFINITION   similar to gb:L11045_rna1 PROTEASOME COMPONENT C13 (HUMAN); mRNA
               sequence.
ACCESSION    AI307267
VERSION      AI307267.1  GI:4002456
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Homnidae; Homo.

Qy 1 CGGTGAAGTCATCTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 110 CGGTGAAGTCATCTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 51

Qy 61 CTGGTGCCCACTAAGTAGAAAAA 85
Db 50 CTGGTGCCCACTAAGTAGAATAA 26

RESULT 8
BF437551/c
LOCUS         BF437551
DEFINITION   7574808.x1 NCI_CGAP_P-28 Homo sapiens cDNA clone IMAGE:3651374 3',
               mRNA sequence.
ACCESSION    BF437551
VERSION      BF437551.1  GI:11449957
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Homnidae; Homo.

REFERENCE    1 (bases 1 to 272)
AUTHORS       NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE         National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL       Unpublished (1997)
COMMENT       Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-remail.nih.gov
               Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
               Emmert-Buck, M.D., Ph.D.
               cDNA Library Preparation: M. Bento Soares, Ph.D.

```

```

REFERENCE      1 (bases 1 to 272)
AUTHORS       NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE         National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL       Unpublished (1997)
COMMENT       Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-remail.nih.gov
               Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
               Emmert-Buck, M.D., Ph.D.
               cDNA Library Preparation: M. Bento Soares, Ph.D.
               DNA Sequencing by: Washington University Genome Sequencing Center
               Clone distribution: NCI-CGAP clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               www-bio.llnl.gov/bbrp/image/image.html
               Seq primer: -40UP from Gibco.
               Location/Qualifiers
               1..272
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:2055283"
                 /tissue_type="2 pooled tumors (clear cell type)"
                 /lab_host="DH10B"
                 /clone_lib="NCI CGAP_Kid12"
                 /note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
                 a modified polylinker; Site 1: Not I; Site 2: Eco RI;
                 Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
                 prepared, and ss circles were made in vitro. Following HAP
                 purification, this DNA was used as tracer in a subtractive
                 hybridization reaction. The driver was PCR-amplified cDNAs
                 from a pool of 5,000 clones made from the same library
                 (cloneIDs 1323912-1325831, 1471368-1472903 and
                 1492104-1493255). Subtraction by Bento Soares and M.
                 Fatima Bonaldo."

ORIGIN
               Query Match      98.1%; Score 83.4; DB 1; Length 272;
               Best Local Similarity 98.8%; Pred. No. 4e-18;
               Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTCATCTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db 112 CGGTGAAGTCATCTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 53

Qy 61 CTGGTGCCCACTAAGTAGAAAAA 85
Db 52 CTGGTGCCCACTAAGTAGAATAA 28

RESULT 9
BF437551/c
LOCUS         BF437551
DEFINITION   7574808.x1 NCI_CGAP_P-28 Homo sapiens cDNA clone IMAGE:3651374 3',
               mRNA sequence.
ACCESSION    BF437551
VERSION      BF437551.1  GI:11449957
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Homnidae; Homo.

REFERENCE    1 (bases 1 to 272)
AUTHORS       NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE         National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL       Unpublished (1997)
COMMENT       Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-remail.nih.gov
               Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
               Emmert-Buck, M.D., Ph.D.
               cDNA Library Preparation: M. Bento Soares, Ph.D.

```


cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov

Seq primer: -40UP from Gibco.

FEATURES

Location/Qualifiers
1..272
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3651374"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pr28"

/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Pr22 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 98.1%; Score 83.4; DB 2; Length 272;
Best Local Similarity 98.8%; Pred. No. 4e-18; Indels 0; Gaps 0;
Matches 84; Conservative 0; Mismatches 1;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 60
|||||
Db 104 CGGTGAAGTGCATCTTCTCGGTGTTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 45

QY 61 CTGGTGCCCACTAAGTAGAATAA 85
|||||

Db 44 CTGGTGCCCACTAAGTAGAATAA 20
|||||

RESULT 9

AI400926/c
LOCUS 273 bp mRNA linear EST 30-MAR-1999
DEFINITION th25e10.x1 NCI CGAP Pr28 Homo sapiens cDNA clone IMAGE:2119338 3', similar to gb:L11045_rnal PROTEASOME COMPONENT C13 (HUMAN); mRNA sequence.

ACCESSION AI400926

VERSION AI400926.1 GI:4244013

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 273)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 348 Std Error: 0.00

Seq primer: -40UP from Gibco.

Location/Qualifiers

1..273

FEATURES

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2119338"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pr28"

/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Pr22 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 98.1%; Score 83.4; DB 1; Length 273;

Best Local Similarity 98.8%; Pred. No. 4e-18; Indels 0; Gaps 0;

Matches 84; Conservative 0; Mismatches 1;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 60
|||||

Db 105 CGGTGAAGTGCATCTTCTCGGTGTTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 46
|||||

QY 61 CTGGTGCCCACTAAGTAGAATAA 85
|||||

Db 45 CTGGTGCCCACTAAGTAGAATAA 21
|||||

RESULT 10

AI669789/c

LOCUS 273 bp mRNA linear EST 15-DEC-1999

DEFINITION tuc1c11.x1 NCI CGAP Pr28 Homo sapiens cDNA clone IMAGE:2252660 3', similar to gb:L11045_rnal PROTEASOME COMPONENT C13 (HUMAN); mRNA sequence.

ACCESSION AI669789

VERSION AI669789.1 GI:4834563

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 273)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 379 Std Error: 0.00

Seq primer: -40UP from Gibco.

Location/Qualifiers

1..273

FEATURES

source

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2252660"

/sex="male"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="NCI CGAP Pr28"


```

QY 61 CTGGGTGCCCACTAAGTAGAAAA 85
|||||
Db 39 CTGGGTGCCCACTAAGTAGAATA 15

RESULT 13
AI417902/c
LOCUS
DEFINITION
t955b11.x1 NCI CGAP Pr28 mRNA linear EST 30-MAR-1999
similar to gb:L11045_rnal PROTEASOME COMPONENT C13 (HUMAN); mRNA
sequence.
ACCESSION
VERSION AI417902.1 GI:4261406
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 304)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 334 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1..304
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2112669"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pr28"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI CGAP Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneIDs
985608-986759, 1101192-1101959, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 98.1%; Score 83.4; DB 1; Length 304;
Best Local Similarity 98.8%; Pred. No. 4.1e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 60
|||||
Db 100 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 41

QY 61 CTGGGTGCCCACTAAGTAGAAAA 85
|||||
Db 40 CTGGGTGCCCACTAAGTAGAATA 16

RESULT 14
AI760920/c
LOCUS
DEFINITION
w170b11.x1 NCI CGAP Kid12 Homo sapiens cDNA clone IMAGE:2398653 3',
similar to gb:L11045_rnal PROTEASOME COMPONENT C13 (HUMAN); mRNA
sequence.
ACCESSION
VERSION AI760920.1 GI:5176587
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 306)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 387 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1..306
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2398653"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid12"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Kid5 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1323912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fatima Bonaldo."

ORIGIN
Query Match 98.1%; Score 83.4; DB 1; Length 306;
Best Local Similarity 98.8%; Pred. No. 4.1e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 60
|||||
Db 112 CGGTGAAGTGCATCTCTCGGTGTTCTCTATTGAAACAGCATTTCCCCAGGGAAGTTT 53

QY 61 CTGGGTGCCCACTAAGTAGAAAA 85
|||||
Db 52 CTGGGTGCCCACTAAGTAGAATA 28

RESULT 15
AA148966/c
LOCUS
DEFINITION
zn99g11.s1 Stratagene colon (#937204) Homo sapiens cDNA clone
IMAGE:566372 3', similar to gb:L11045_rnal PROTEASOME COMPONENT C13
(HUMAN); mRNA sequence.
ACCESSION
VERSION AA148966.1 GI:1719064
KEYWORDS EST.
SOURCE Homo sapiens (human)

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 309)
Hallier, L., Lennon, G., Becker, M., Bonaudo, M.F., Chiapelli, B.,
Chisoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaskie, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.
and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
PUBMED 8899549
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

FEATURES
source 1..309
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:4590825"
 /db_xref="taxon:9606"
 /clone="IMAGE:566372"
 /tissue="tumor"
 /cell_line="T84 carcinoma cell line"
 /lab_host="SOLR cells (kanamycin resistant)"
 /clone_lib="Stratagene colon (#937204)"
 /notes="Organ: colon; Vector: pBluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. T-84 colonic epithelial cell line. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
CTCGAGTATTTTATTTTATTTT 3'"

ORIGIN

Query Match 98.1%; Score 83.4; DB 1; Length 309;
Best Local Similarity 98.8%; Pred. No. 4.1e-18;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGCAACAGCATTTCCCCAGGGAGTTT 60
 |||||
Db 103 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGCAACAGCATTTCCCCAGGGAGTTT 44
 |||||

Qy 61 CTGGGTGCCCCACTAAGTAGAAAAA 85
 |||||
Db 43 CTGGGTGCCCCACTAAGTAGAATAA 19
 |||||

Search completed: December 7, 2005, 09:56:06
Job time : 973.311 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	93.4	25.2	79835	3	US-09-949-016-12456		Sequence 12456, A
2	93.4	25.2	79835	3	US-09-949-016-16121		Sequence 16121, A
C 3	90	24.3	9527	3	US-09-949-016-13779		Sequence 13779, A
	89.4	24.2	20481	3	US-09-949-016-12093		Sequence 12093, A
5	89.4	24.2	20482	3	US-09-949-016-13660		Sequence 13660, A
6	88.6	23.9	601	3	US-09-949-016-143891		Sequence 143891, A
7	88.6	23.9	601	3	US-09-949-016-143892		Sequence 143892, A
8	88.6	23.9	601	3	US-09-949-016-143893		Sequence 143893, A
9	88.6	23.9	601	3	US-09-949-016-143894		Sequence 143894, A
10	88.6	23.9	601	3	US-09-949-016-143895		Sequence 143895, A
11	88.6	23.9	15330	3	US-09-949-016-15808		Sequence 15808, A
12	87.6	23.7	222452	3	US-09-949-016-12968		Sequence 12968, A
C 13	87	23.5	152070	3	US-09-949-016-15402		Sequence 15402, A
	86.6	23.4	31444	3	US-09-949-016-17532		Sequence 17532, A
15	86.2	23.3	25260	3	US-09-949-016-11985		Sequence 11985, A
16	86.2	23.3	25260	3	US-09-949-016-12907		Sequence 12907, A
C 17	86	23.2	63783	3	US-09-949-016-13576		Sequence 13576, A
	85.8	23.2	601	3	US-09-949-016-64807		Sequence 64807, A
18	85.8	23.2	601	3	US-09-949-016-13372		Sequence 13372, A
19	85.8	23.2	21590	3	US-09-949-016-13732		Sequence 13732, A
20	85.8	23.2	21590	3	US-09-949-016-13732		Sequence 13732, A
21	85.6	23.1	28819	3	US-09-949-016-15806		Sequence 15806, A
22	85.4	23.1	61124	3	US-09-949-016-11914		Sequence 11914, A
23	85.4	23.1	61140	3	US-09-949-016-15771		Sequence 15771, A
24	84.8	22.9	601	3	US-09-949-016-42374		Sequence 42374, A


```
Query Match      24.2%; Score 89.4; DB 3; Length 20482;
Best Local Similarity 75.5%; Pred. No. 1.5e-19;
Matches 111; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 70 AGGTCATCATCTGAGGTTTCATCAACACATGGAAGCCATGCTAGTATTGTCCTTCGAT 129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12360 AGGCCACATGCTGAGGTTTCATGAAGCACTGGAACCATGGTAGTTGTTCTTGAT 12419

Qy 130 AGGTATGTCACCTCAACACCAATGCTCTCTGCTGATGAGAAATGGGCTGCCCGAT 189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12420 AAGTATGTCGGCTCAACACCAAGTGCCTCTCTGCTGATGAGAAATGGATGGCACTT 12479

Qy 190 TCACAGACTTCTTGATCTTTCTCTCAGC 216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12480 GCCCTGAAAGATGTCAGCTGCCCAGC 12506

RESULT 6
US-09-949-016-143891/c
; Sequence 143891, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 143891
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-143891

Query Match      23.9%; Score 88.6; DB 3; Length 601;
Best Local Similarity 72.3%; Pred. No. 4.9e-20;
Matches 115; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 212 TCAGCATATGATCTCTCATATGAACCTCTCTCTGCTAGGAGTGCGAGCTTACATCG 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 556 TCTCCTTTTCTCCACTTCTGACAGATCTAGTCTTTGGCCAGGCGTGGTGGCTCACATCT 497

Qy 272 CTAATCCCGCAATACAGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 496 ATAATCTAGCACTTTGGAGGCTCAGGAGGAGGATCATTGAAATCCAGGAGTTTGAGA 437

Qy 332 CCAGGCTGGCAATACAGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 436 CCAGTCTGGCAACACAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 398

RESULT 7
US-09-949-016-143892/c
; Sequence 143892, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
```

```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 143892
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-143892

Query Match      23.9%; Score 88.6; DB 3; Length 601;
Best Local Similarity 72.3%; Pred. No. 4.9e-20;
Matches 115; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 212 TCAGCATATGATCTCTCATATGAACCTCTCTCTGCTAGGAGTGCGAGCTTACATCG 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 544 TCTCCTTTTCTCCACTTCTGACAGATCTTAGTCTTTGGCCAGGCGTGGTGGCTCACATCT 485

Qy 272 CTAATCCCGCAATACAGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 484 ATAATCTAGCACTTTGGAGGCTCAGGAGGAGGATCATTGAAATCCAGGAGTTTGAGA 425

Qy 332 CCAGGCTGGCAATACAGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 424 CCAGTCTGGCAACACAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 386

RESULT 8
US-09-949-016-143893/c
; Sequence 143893, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 143893
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-143893

Query Match      23.9%; Score 88.6; DB 3; Length 601;
Best Local Similarity 72.3%; Pred. No. 4.9e-20;
Matches 115; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 212 TCAGCATATGATCTCTCATATGAACCTCTCTCTGCTAGGAGTGCGAGCTTACATCG 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 512 TCTCCTTTTCTCCACTTCTGACAGATCTTAGTCTTTGGCCAGGCGTGGTGGCTCACATCT 453

Qy 272 CTAATCCCGCAATACAGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 452 ATAATCTAGCACTTTGGAGGCTCAGGAGGAGGATCATTGAAATCCAGGAGTTTGAGA 393

Qy 332 CCAGGCTGGCAATACAGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 392 CCAGTCTGGCAACACAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 354

RESULT 9
US-09-949-016-143894/c
; Sequence 143894, Application US/09949016
```


Db 23428 TCAAAATTATGCTTCTCGCTGGGCGTGGTGGCTCATACTTCTTATCCAGCACTTTGGGAG 23487
Qy 293 GCTGAGGCAGGAGATTACTGGAGGCCAGGAGTTTGAGACCAGGCTGGGCAATACAGCGA 352
Db 23488 GCGAGGCAGGAGATCACTTGAGGCTAGGAGTTTGAGACCAGGCTGGGCAACACAGCAA 23547
Qy 353 GACTCTCTCTC 363
Db 23548 GACCTATCTC 23558

Search completed: December 6, 2005, 11:47:01
Job time : 164.493 secs

Result No.	Score	Query #		Length	DB	ID	Description
		Match					
1	368.4	99.6	202381	8	AC091167	AC091167 Homo sapi	
2	95.4	25.8	116061	8	AL732324	AL732324 Human DNA	
3	94.4	25.5	163863	8	AC096971	AC096971 Homo sapi	
4	93.4	25.2	9933	8	HSMTGR1E10	AF076458 Homo sapi	
5	93.4	25.2	81944	8	HSDB63M2	AL121906 Human DNA	
c	93						
6	92.2	24.9	156912	8	AL160176	AL160176 Human DNA	
7	92.2	24.9	153922	8	AC011921	AC011921 Homo sapi	
8	92.2	24.9	174711	14	AC025515	AC025515 Homo sapi	
9	91.8	24.8	99163	8	AL355602	AL355602 Human DNA	
10	91.8	24.8	165079	14	AC024517	AC024517 Homo sapi	
c	91.8						
11	91.8	24.8	167300	14	AC021394	AC021394 Homo sapi	
12	91.4	24.7	182509	8	AC112211	AC112211 Homo sapi	
c	91.2	24.6	101769	14	AC093303	AC093303 Homo sapi	
13	91.2	24.6	128540	14	AC025473	AC025473 Homo sapi	
c	91.2						
14	91.2	24.6	141633	8	AC073542	AC073542 Homo sapi	
15	91.2	24.6	141633	8	AC073542	AC073542 Homo sapi	
16	91.2	24.6	178042	14	AC036168	AC036168 Homo sapi	
17	91.2	24.6	182892	8	AC034244	AC034244 Homo sapi	
c	91.2						
18	91.2	24.6	195068	8	AC012100	AC012100 Homo sapi	


```

repeat_region complement(8454..8748)
repeat_region /rpt family="AluSx"
repeat_region complement(8749..8921)
repeat_region /rpt family="MER82"
repeat_region 9286..9576
repeat_region /rpt family="AluJo"
repeat_region 9586..9884
repeat_region /rpt family="AluSx"
repeat_region 9971..10186
repeat_region /rpt family="L1PA3"
repeat_region complement(10138..10482)
repeat_region /rpt family="MER82"
repeat_region complement(10581..10785)
repeat_region /rpt family="L1ME"
repeat_region complement(10790..11028)

Query Match 99.6%; Score 368.4; DB 8; Length 202381;
Best Local Similarity 99.7%; Pred. No. 3.8e-108;
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCGTTTCAGCCACAGCAATGTATAGTGTCTACAACTCTTATACAGGGATACAGCAATG 60
Db 90994 CCGTTTCAGCCACAGCAATGTATAGTGTCTACAACTCTTATACAGGGATACAGCAATG 91053

Qy 61 TGAACACAGAGGTATGATCTGAGTTTCATCAAAACACTGGAAGCCATGGTAGTTATT 120
Db 91054 TGAACACAGAGGTATGATCTGAGTTTCATCAAAACACTGGAAGCCATGGTAGTTATT 91113

Qy 121 GTCTTTGATAGTATGTCAACTCAACACCAATGCTCTCTCTGCTGCTGATGAGAAATGGG 180
Db 91114 GTCTTTGATAGTATGTCAACTCAACACCAATGCTCTCTCTGCTGCTGATGAGAAATGGG 91173

Qy 181 CTGCCGATTCACAGACTTCTTGATCTTCTCTCAGCATATGATATCTCTCATATCAAACTC 240
Db 91174 CTGCCGATTCACAGACTTCTTGATCTTCTCTCAGCATATGATATCTCTCATATCAAACTC 91233

Qy 241 TCCTCTCGGTAGAGTGGCAGCTTACATCTCTATATCCACACCTTGGGAAGCTGAGGC 300
Db 91234 TCCTCTCGGTAGAGTGGCAGCTTACATCTCTATATCCACACCTTGGGAAGCTGAGGC 91293

Qy 301 AGGAGGATTCTGAGGCGCAGGTTTGAGACCCAGGCTGGCAATACAGCGAGCTCTCT 360
Db 91294 AGGAGGATTCTGAGGCGCAGGTTTGAGACCCAGGCTGGCAATACAGCGAGCTCTCT 91353

Qy 361 CTCCTAAAAA 370
Db 91354 CTCCTAAAAA 91363

RESULT 2
AL732324 116061 bp DNA linear PRI 18-MAY-2005
LOCUS Human DNA sequence from clone RP13-470G3 on chromosome X Contains
DEFINITION the 5' end of the STAR88 gene for START domain containing 8, a
ribosomal protein L31 (RPL31) pseudogene and a cytochrome c oxidase
subunit Vtc (COX6C) pseudogene, complete sequence.
ACCESSION AL732324
VERSION AL732324.6 GI:21667382
KEYWORDS HTG; COX6C; RPL31; STAR88.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 116061)
Whitehead,S.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Jul 2, 2002 this sequence version replaced gi:21531526.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:

```

```

Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
RP13-470G3 is from the library RPCI-13.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
FEATURES             location/Qualifiers
source               1..116061
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /clone="RP13-470G3"
                    /clone_1lb="RPCI-13.2"
                    37720..38094
gene                /locus_tag="RP13-470G3.1-001"
                    /pseudo
                    37720..38094
CDS                 /locus_tag="RP13-470G3.1-001"
                    /note="match: proteins: AAH50113 CAA48925 O18602 O65071
                    P12947 Q9VY77 Q9BV33 Q9CY93 Q9GN74 Q9GP16 Q9IA76 Q9SL77
                    Q9U332 Q9V597 Q9XGL4"
                    /pseudo
                    /codon_start=1
                    /product="ribosomal protein L31 (RPL31) pseudogene"
                    complement(82825..82830)
polyA_signal        complement(82947..83148)
gene                /locus_tag="RP13-470G3.2-001"
                    /pseudo
                    complement(82947..83148)
CDS                 /locus_tag="RP13-470G3.2-001"
                    /note="match: proteins: P04038 P11950 P11951 Q9CPQ1"
                    /pseudo
                    /codon_start=1
                    /product="cytochrome c oxidase subunit Vtc (COX6C)
                    pseudogene"
gene                join(85287..85548,93740..94172)
                    /gene="STAR88"
                    /locus_tag="RP13-469L5.1-002"
mRNA                join(85287..85548,93740..94172)
                    /gene="STAR88"
                    /locus_tag="RP13-469L5.1-002"
                    /product="START domain containing 8"
                    /note="match: ESTs: B114834.1"
gene                join(85290..85548,103120..103153,AL3360076.9:36474..36545,
                    AL360076.9:39928..39991,AL360076.9:40772..42189,
                    AL360076.9:42785..42958,AL360076.9:43824..43983,
                    AL360076.9:44484..44682,AL360076.9:45096..45306,
                    AL360076.9:45607..45721,AL360076.9:45982..46206,
                    AL360076.9:47186..47403,AL360076.9:47505..47681,
                    AL360076.9:47865..479402)
                    /gene="STAR88"
                    /locus_tag="RP13-469L5.1-001"
mRNA                join(85290..85548,103120..103153,AL3360076.9:36474..36545,
                    AL360076.9:39928..39991,AL360076.9:40772..42189,
                    AL360076.9:42785..42958,AL360076.9:43824..43983,
                    AL360076.9:44484..44682,AL360076.9:45096..45306,
                    AL360076.9:45607..45721,AL360076.9:45982..46206,
                    AL360076.9:47186..47403,AL360076.9:47505..47681,
                    AL360076.9:47865..479402)
                    /gene="STAR88"
                    /locus_tag="RP13-469L5.1-001"

```


952	947	1807	1801	3673	3751
4432	4431	78	<800	1705	1698
419	<800	6598	6646	2125	2048
7007	7341	3832	3739	1223	1207
3030	2995	511	<800	4770	4682
1158	1135	2564	2516	6497	6529
455	<800	457	<800	2009	2048
5803	5650	1272	1242	4725	4682
4738	4591	2179	2205	690	<800
7220	7341	799	822	13586	13272
1436	1372	189	<800	4512	4682
10664	10373	258	<800	3252	3173
3133	3141	1085	1060	1909	1908
264	<800	4316	4265	6651	6529
40	<800	10260	10186	1086	1070
687	<800	7100	7137	322	<800
6385	6420	1203	1242	521	<800
958	947	163	<800	90	<800
887	947	3660	3633	611	<800
821	819	10512	10578	17096	17342
4910	4943	3038	3051	559	<800
1754	1702	5181	5182	9637	9470
1547	1547	382	<800	4351	4682
536	<800	1372	1338	2432	2440
37	<800	2042	1994	1736	1698
4637	4591	8220	8316	1556	1538
1602	1547	1068	1060	2585	2594
1860	1830	7883	7986	1963	1908
829	819	139	<800	1746	1698
2153	2139	608	<800	4639	4682
689	<800	2276	2205	1488	1461
7446	7341	1681	1721	1897	1908
511	<800	6387	6435	4759	4682
656	<800	4559	4500	380	<800
993	947	1832	1801	1687	1698
174	<800	5347	5362	584	<800
2644	2648	11879	11556	1602	1698

Query Match 25.5%; Score 94.4; DB 8; Length 163863;
Best Local Similarity 80.9%; Pred. No. 4.6e-19;
Matches 110; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 66 CCAGAGTCTATGTCATCTGAGTTTCATCAAAACACCTGGAAGCCATGCTAGTATTGTCCT 125
Db 139652 CCAAGTCCATGTCATCTGAGTTTGTAAAGCACCTGGAAGCCATGCTAGTATTGTCCT 139711

Qy 126 TGATAGTATGTCAACCTCAACCAATGCTCTCTGGCTGATGGAGAAATGGGCTGCC 185
Db 139712 TGACAGGTATGTCAACCTTAACCAACGCTCTTCTCTGGCTGATGGAGAAATGGATGCAG 139771

Qy 186 CGATTCCACACTTCT 201
Db 139772 CCAAGATCAGAAATCT 139787

RESULT 4
HSMGTGR1E10 933 bp DNA linear PRI 18-JAN-2000
LOCUS Homo sapiens MTG8-like protein (MTG8) exon 10, complete sequence.
DEFINITION AF076458
ACCESSION AF076458.1 GI:5256960
VERSION 1
KEYWORDS 10 of 13
SEGMENT Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 933)
AUTHORS Kitabayashi, I., Ida, K., Morohoshi, F., Yokoyama, A., Mitsuhashi, N., Shimizu, K., Nomura, N., Hayashi, Y. and Ohki, M.
TITLE The AML1-MTG8 leukemic fusion protein forms a complex with a novel member of the MTG8(ETO/CDR) family, MTGR1
JOURNAL Mol. Cell. Biol. 18 (2), 846-858 (1998)
PUBMED 9447981
REFERENCE 2 (bases 1 to 933)
AUTHORS Morohoshi, F., Mitani, S., Mitsuhashi, N., Kitabayashi, I., Takahashi, E., Suzuki, M., Munakata, N. and Ohki, M.
TITLE Structure and expression pattern of a human MTG8/ETO family gene, MTGR1
JOURNAL Gene 241 (2), 287-295 (2000)
PUBMED 10675041
REFERENCE 3 (bases 1 to 933)
AUTHORS Morohoshi, F., Mitani, S., Mitsuhashi, N., Kitabayashi, I., Takahashi, I., Suzuki, M., Munakata, N. and Ohki, M.
TITLE Direct Submission
JOURNAL Submitted (07-JUL-1998) Radiobiology, National Cancer Center Research Institute, Tsukiji, 5-1-1, Chuo-ku, Tokyo 104-0045, Japan
FEATURES Location/Qualifiers
source 1. .933
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
394. .589
/gene="MTGR1"
/number=10
exon

```
ORIGIN
Query Match      25.2%; Score 93.4; DB 8; Length 933;
Best Local Similarity 80.7%; Pred. No. 6.9e-19;
Matches 109; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 50 ATACAGCAATGTGAACACGAGGTCATGCTCTGAGGTTTCATCAAAACACTGGAAGCCA 109
    |||
Db 115 ATAGTAGACTGTGGCCAGAGGCCATGATCTGAGGTTTGTCCAAACACTGGAAGCCC 174
    |||

QY 110 TGGTAGTATTGCTCTTGTAGTAGTATGTCACACCTCAACACCAATGCTCTTCCTGGCTGAT 169
    |||
Db 175 TGGTAGTATTGCTCTTGTAGTAGTATGTCAGCCTCAACACCAATGCTCTTCCTAGCTGAT 234
    |||

QY 170 GGAGAAATGGCGTGC 184
    |||
Db 235 GGAGAAACACACAGC 249
    |||

RESULT 5
HSDJ63M2      81914 bp      DNA      linear      PRI 18-MAY-2005
LOCUS         Human DNA sequence from clone RP1-63M2 on chromosome 20 Contains
DEFINITION    the 3' end of the CBFA2T2 gene for runt domain core-binding factor
              alpha subunit 2, the C20orf134 gene, the E2F1 gene for E2F
              transcription factor 1, the APB2BP gene for amyloid beta (A4)
              precursor protein-binding family A, member 2 protein, the C20orf144
              gene and six CpG islands, complete sequence.
              AL121906
ACCESSION     AL121906.18 GI:8574371
VERSION       Mammalia; Euarchontoglires; Primates; Catarrhini;
KEYWORDS      E2F1; MTGR1; transcription factor.
SOURCE        Homo sapiens
ORGANISM      Homo sapiens
              Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominoidea; Homo.
              1 (bases 1 to 81914)
REFERENCE     Bird.C.
AUTHORS       Direct Submission
TITLE         Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL       Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
              Clone requests: clonerequests@sanger.ac.uk
              On Jun 20, 2000 this sequence version replaced gi:8574056.
              The following abbreviations are used to associate primary accession
              numbers given in the feature table with their source databases:
              En:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
              on the WORMPEP database can be found at
              http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
              was generated from part of bacterial clone contigs of human
              chromosome 20, constructed by the Sanger Centre Chromosome 20
              Mapping Group. Further information can be found at
              http://www.sanger.ac.uk/HGP/Chr20
              RP1-63M2 is from the library RPC1-1 constructed by the group of
              Pletier de Jong. For further details see
              http://www.chori.org/bacpac/home.htm
              VECTOR: pCYPAC2
              ----- Genome Center
              Center: Wellcome Trust Sanger Institute
              Center code: SC
              Web site: http://www.sanger.ac.uk
              Contact: vegas@sanger.ac.uk
              -----
              This sequence was finished as follows unless otherwise noted: all
              regions were either double-stranded or sequenced with an alternate
              chemistry or covered by high quality data (i.e., phred quality >=
              30); an attempt was made to resolve all sequencing problems, such
              as compressions and repeats; all regions were covered by at least
              one subclone; and the assembly was confirmed by restriction digest,
              except on the rare occasion of the clone being a YAC.
              Location/Qualifiers
                1. .81914
                  /organism="Homo sapiens"

FEATURES
Source
```

```
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP1-63M2"
/clone_lib="RPC1-1"
100
misc_feature
100
note="Clone right end: RP5-1137F22"
join(AL034421.7:88709..88750,AL034421.7:92259..92440,
210..291,1201..1454,4787..4872,6156..6351,13077..13099)
/locus_tag="RP5-1137F22.1-006"
/gene="CBFA2T2"
join(AL034421.7:88709..88750,AL034421.7:92259..92440,
210..291,1201..1454,4787..4872,6156..6351,13077..13099)
/gene="CBFA2T2"
/locus_tag="RP5-1137F22.1-006"
note="match: ESTs: BQ434235.1"
join(AL034421.7:31478..31818,AL034421.7:76100..76243,
AL034421.7:80238..80479,AL034421.7:88661..88750,
AL034421.7:92259..92440,210..391)
/gene="CBFA2T2"
/locus_tag="RP5-1137F22.1-005"
note="match: ESTs: AA635096.1 AU100534.1 BF679319.1
BM725921.1"
match: cDNAs: BC015066.1"
join(AL034421.7:76126..76243,AL034421.7:80238..80479,
AL034421.7:88661..88750,AL034421.7:92259..92440,210..369)
/gene="CBFA2T2"
/locus_tag="RP5-1137F22.1-005"
/standard_name="OTTHUMP0000030655"
/codon_start=1
/protein_id="CAI22130.1"
/db_xref="GI:56203747"
/translation="MPGSPVEVKIQSRSSPPTMPPLPPINPGGPRPVSTPTALSNGL
NHSPTLNGAPSPORFSGPASSTSSALTNOQLPATCAGQSLKURKELTITQQFGN
DISPRIGEKVRLVLALVNSTVTEEFHCKLQEAATNFFPLRPFVIFPKLALPQREL
LHCARAAKQTSQYLAQHEHLLNTSIASPADSELLMEVHGNGKFPSPESCIIHLKS
MGVASRHEYSGTLOWPAHPVNSTLENLWDCSRLSRSTVLPFP"
join(AL034421.7:43164..43405,AL034421.7:76100..76243,
AL034421.7:80238..80479,AL034421.7:88661..88750,
AL034421.7:92259..92440,1201..1454,4787..4872,6156..6351,
13077..13145,16778..16968,20784..26473)
/gene="CBFA2T2"
/locus_tag="RP5-1137F22.1-003"
join(AL034421.7:31524..31818,AL034421.7:43118..43405,
AL034421.7:70974..71049,AL034421.7:76100..76243,
AL034421.7:80238..80479,AL034421.7:88661..88750,
AL034421.7:92259..92440,1201..1454,4787..4872,6156..6351,
13077..13145,16778..16968,20784..26473)
/gene="CBFA2T2"
/locus_tag="RP5-1137F22.1-004"
join(AL034421.7:31524..31818,AL034421.7:43118..43405,
AL034421.7:70974..71049,AL034421.7:76100..76243,
AL034421.7:80238..80479,AL034421.7:88661..88750,
AL034421.7:92259..92440,1201..1454,4787..4872,6156..6351,
13077..13145,16778..16968,20784..26473)
/gene="CBFA2T2"
/locus_tag="RP5-1137F22.1-003"
note="match: ESTs: AA459492.1 AU017082.1 AW060163.1
AW952627.1 BE741379.1 BE743916.1 BQ003927.1
BU737758.1 CA307223.1"
match: cDNAs: AF069747.1"
join(AL034421.7:43164..43405,AL034421.7:76100..76243,
AL034421.7:80238..80479,AL034421.7:88661..88750,
AL034421.7:92259..92440,1201..1454,4787..4872,6156..6351,
13077..13145,16778..16968,20784..26473)
/gene="CBFA2T2"
/locus_tag="RP5-1137F22.1-003"
note="match: ESTs: AJ948413.1 AW060163.1 BE741379.1
```


misc_feature	db_xref="RZPD:RPCI704F20579"			
	db_xref="taxon:9606"			
	chromosome="20"			
	clone="RP4-579F20"			
	clone_lib="RPC1-4"			
	100			
	note="Clone right_end: RP4-800J21"			
	complement(20584)			
	complement(20585..21785)			
	gene="HMG1L1"			
polyA_site	locus_tag="RP4-579F20.2-001"			
	complement(20585..21785)			
	gene="HMG1L1"			
	locus_tag="RP4-579F20.2-001"			
	product="match: ESTs: AA045606.1 AA050755.1 AA054783.1 AA093946.1 AA101746.1 AA118599.1 AA150131.1 AA196344.1 AA206991.1 AA209397.1 AA220281.1 AA220725.1 AA220730.1 AA247271.1 AA275567.1 AA311991.1 AA313759.1 AA315282.1 AA316546.1 AA319424.1 AA330028.1 AA335709.1 AA360132.1 AA360952.1 AA370554.1 AA385117.1 AA445091.1 AA459233.1 AA507781.1 AA524370.1 AA541647.1 AA546572.1 AA557452.1 AA563485.1 AA576844.1 AA600212.1 AA607172.1 AA607305.1 AA641615.1 AA641999.1 AA715713.1 AA745955.1 AA795527.1 AA797955.1 AA798375.1 AA845496.1 AA861823.1 AA876587.1 AA994792.1 AA998198.1 AA104961.1 AA1041775.1 AA105974.1 AA106499.1 AA1077928.1 AA109696.1 AA1127224.1 AA1127978.1 AA151854.1 AA160153.1 AA1186025.1 AA190052.1 AA1280190.1 AA134636.1 AA1348036.1 AA1351048.1 AA1368179.1 AA1535682.1 AA153705.1 AA1535744.1 AA1590841.1 AA1654646.1 AA1734238.1 AA1734851.1 AA1749800.1 AA1788272.1 AA1811093.1 AA1819512.1 AA1829489.1 AA1889428.1 AA1927573.1 AA1935842.1 AA1956408.1 AA1963188.1 AA1998803 AA1047978 AA048263 AA050935 AA011543 AA014204 AA096099 AV129402 AV134319 AV171164 AV255394 AV301413 AA004848.1 AA016332.1 AA020136.1 AA023735.1 AA026237.1 AA130136.1 AA142663.1 AA151090.1 AA160834.1 AA250057.1 AA258476.1 AA265014.1 AA270920.1 AA314143.1 AA364611.1 AA364933.1 AA364962.1 AA364966.1 AA403673.1 AA464310.1 AA467569.1 AA475171.1 AA476025.1 AA537381.1 AA543443.1 AA544911.1 AA545883.1 AA547929.1 AA549623.1 AA553354.1 AA554198.1 AA555741.1 AA556397.1 AA556870.1 AA557546.1 AA575412.1 AA575638.1 AA610999.1 D57151 P22618 H50065.1 H96457.1 M62207.1 M84830.1 T28966.1 W07682.1 W32423.1 W58023.1 W85536.1			
	match: CDNA8: AB037850 AB037864 AF010465 AF033116 AF049239 AF049240 AF056322 AF099674 AF169946 AF169947 AF169948 AF201422 AF234161 AK001898 D10202 D63874 E03879 I07333 L07334 L13804 L38477 L39018 M21683 M26110 M64986 M76674 M80436 M88749 S26224 U00431 U01343 U01345 U01348 U05593 U10925 U21933 U88622 X12597 X12796 X63463 Y00463 Y14591 Y17153 Y17968 Z11997"			
	complement(20617)			
	gene="HMG1L1"			
	locus_tag="RP4-579F20.2-001"			
	complement(21047..21682)			
gene	locus_tag="RP4-579F20.2-001"			
	standard_name="OTTHUMP00000031372"			
	note="match: proteins: P07155 P09429 P10103 P12682 Q91596 Q9YH06"			
	codon_start=1			
	protein_id="CAB92731.1"			
	db_xref="GI:8217648"			
	db_xref="GOA:Q9NQJ4"			
	db_xref="HSSP:1HMF"			
	db_xref="InterPro:IPR000135"			
	db_xref="InterPro:IPR000910"			
polyA_signal	db_xref="UniProt/TREMBL:Q9NQJ4"			
	translation="MGKGDPKPRGMSSVAFVQICREHKKKHSDASVNFSEFNK CSRWKMTSAKGGEDMAKDKTYERQMKTYPKGETKKKKFDNAPKPPPSAF FLPCSEYHPKIGKEHPLGLSIGDVAKLGLGEMNNNTAADKKQPEKKKALKRYEKDIA AYQAKGKPEAAKGVVKAESKKKKKEEEDDEDEDEDEDEDEDEDEDDDDDE"			
	join(complement(AL035541.15:871..972), complement(56318..56872), complement(55723..55933),			
	complement(56318..56872), complement(55723..55933),			
	complement(56318..56872), complement(55723..55933),			
	complement(56318..56872), complement(55723..55933),			
	complement(56318..56872), complement(55723..55933),			
	complement(56318..56872), complement(55723..55933),			
	complement(56318..56872), complement(55723..55933),			
	complement(56318..56872), complement(55723..55933),			
misc_feature	locus_tag="RP4-579F20.1-001"			
	product="CCCTC-binding factor (zinc finger protein)-like"			
	note="match: ESTs: AL732016 AL732017 AL732020 AL844280 AL844303 AL844319 AL844411 AL844477 BG722482 match: CDNA8: AF133731 AF336042 AK001596 AK001744 U25435 U51037 Z22605"			
	complement(29827)			
	gene="CTCF1"			
	locus_tag="RP4-579F20.1-001"			
	complement(29843..29848)			
	gene="CTCF1"			
	locus_tag="RP4-579F20.1-001"			
	product="CCCTC-binding factor (zinc finger protein)-like"			
CDS	note="match: proteins: O14709 O70162 O95015 O97600 P10078 P28698 P49711 P56270 P56671 P97365 Q02975 Q08705 Q61164 Q61751 Q61898 Q9Y2P0"			
	codon_start=1			
	product="CCCTC-binding factor (zinc finger protein)-like"			
	protein_id="CAI42956.1"			
	db_xref="GI:57208876"			
	db_xref="GOA:Q5JU04"			
	db_xref="InterPro:IPR007087"			
	translation="MAATEISVLSEQFTIKLELMPEKGLKEEKDGVCREKDRHSP SLEAEVTSQAFQDSVLEEEVELVLAPESESKYILTLQTVHTSPAEVLQDMSLLSI QQEGVQVVQPGPGLLWLEGGPQSLQCVVAISIQQLYSFQENEVLPQHALENV MVASNDQVLAETITGLILEEEOEKNQLLAERTKEQLFFVETMSGDSDEIVLT VNSNVQEODQTAGDAEAKASTNQRTKGAKGTFCDCVMFTSGRMSFNHMH KTHTEKPHLCHLCTFTVTLRNHNVTHTGTRFYKNCDCNMARFVTSGLVYRHRY KTHTEKPKCSCKTVASVEASKLKRVRSHGTGERPPQCCQSVASRDTYKLRHMRTH SGEKPYECHICTRTFTQSGTMKIHILQKHGENVKYQCPCATIIARKSDLRVHMRL HAYSAAELKRCYCSAVHERYALIQHKTKEKPEKCKHCKSVACKQERHMTAHITH TGEKPTLCSNCKFRQKLLNNAHFKYHDANFIPVYKSCGCKGKGFERNINURHSE GQFFQEMFPVACRETTARVKEEDVGEVTCMLLNTMDK"			
	complement(join(51153..51546,51862..52032))			
	gene="CTCF1"			
mRNA	locus_tag="RP4-579F20.1-002"			
	complement(join(51153..51546,51862..52032))			
	gene="CTCF1"			
	locus_tag="RP4-579F20.1-002"			
	product="CCCTC-binding factor (zinc finger protein)-like"			
	note="match: ESTs: AW243377"			
	complement(51153)			
	gene="CTCF1"			
	locus_tag="RP4-579F20.1-002"			
	complement(51171..51176)			
polyA_signal	locus_tag="RP4-579F20.1-002"			
	gene="CTCF1"			
	locus_tag="RP4-579F20.1-002"			
	note="Clone left_end: RP4-718J7"			
	25.1%; Score 93; DB 8; Length 56912;			
	Best Local Similarity 74.5%; Pred. No. 1.2e-18;			
	ORIGIN			

Matches 117; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Oy 207 TTCTCTCAGCATATGATCTCTCATATGAAACTCTCTCTCGGTAGGAGTGGCAGCCTA 266
 |||||
 Db 48140 TTGCTTCGTAGATTTTCAGGTATTTAAACTCACTCTCGGCAGGATGGTGGCTCA 48081
 |||||
 Oy 267 CATCGTAAATCCAGCACCTTTGGAGCTGAGGAGGAGGATTAATCTGAGGCCAGGAGGTT 326
 |||||
 Db 48080 AACCTGTAAATCCAGCACCTTTGGAGGCTGAGGAGGAGGATTCCTTGAGTCCAGGAGTT 48021
 |||||
 Oy 327 TGAGACCAAGGCTGGCATACAGGAGGAGCTCTCTC 363
 |||||
 Db 48020 TGAGACCAAGGCTGGTCAACATAGGAGACCTGTCTC 47984
 |||||

RESULT 7
 AC011921
 LOCUS AC011921 153922 bp DNA linear PRI 24-MAR-2002
 DEFINITION Homo sapiens chromosome 17, clone RP11-15E18, complete sequence.
 AC011921
 VERSION AC011921.14 GI:19698712
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 153922)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 17, clone RP11-15E18
 Unpublished
 2 (bases 1 to 153922)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
 Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhalter,B.,
 Brown,A., Castle,A., Colangelo,M., Collins,S., Collamore,A.,
 Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donegan,L., Doyle,M.,
 Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
 Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
 Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
 McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
 Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
 Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (16-OCT-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 153922)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A.,
 Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Fato,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kelle,C., Lacroque,K., Lamazares,R.,
 Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
 Mihova,C., Norman,C.H., O'Connor,T., Naylor,J., Nguyen,C., Nicol,R.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Rietta,R., Riback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigliolo,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS

JOURNAL

COMMENT

Submitted (24-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 24, 2002 this sequence version replaced gi:17977432.
 All repeats were identified using RepeatMasker:
 Sait, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L3450
 Center clone name: 15_E_18

FEATURES

source

Location/Qualifiers
 1. 153922
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17"
 /clone="RP11-15E18"
 /clone_lib="RPCI-11 Human Male BAC"
 630..938
 /rpt_family="AluSx"
 963..1713
 /rpt_family="L1MC4a"
 1711..1778
 /rpt_family="L1M4"
 1779..2089
 /rpt_family="AluSx"
 2090..2274
 /rpt_family="L1M4"
 complement(2275..2565)
 /rpt_family="AluSx"
 2566..2622
 /rpt_family="L1M4"
 2623..2926
 /rpt_family="AluSx"
 2927..2955
 /rpt_family="L1M4"
 2956..3270
 /rpt_family="AluJb"
 3271..3466
 /rpt_family="L1M4"
 3899..3931
 /rpt_family="AT rich"
 complement(3935..4247)
 /rpt_family="AluSx"
 5108..5408
 /rpt_family="AluJb"
 complement(5484..5541)
 /rpt_family="MIR"
 5616..5637
 /rpt_family="AT rich"
 complement(5670..5953)
 /rpt_family="AluSx"
 complement(6006..6304)
 /rpt_family="AluJb"
 6315..6338
 /rpt_family="AT rich"
 complement(6347..6631)
 /rpt_family="AluJb"
 7751..7852
 /rpt_family="HAL1b"
 complement(8253..8546)
 /rpt_family="AluSx"
 9020..9322
 /rpt_family="L2"
 9323..9614
 /rpt_family="AluSx"
 9615..9726

TITLE

O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, N., Travis, N., Trigliio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (24-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 24, 2002 this sequence version replaced gi:14210624.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L8265
Center clone name: 634_F_5

----- Summary Statistics
Sequencing vector: M13; M77815; 39% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 168986 bases at least Q40
Consensus quality: 171447 bases at least Q30
Consensus quality: 172390 bases at least Q20
Insert size: 182000; agarose-fp
Insert size: 173111; sum-of-contigs
Quality coverage: 11.1 in Q20 bases; agarose-fp
Quality coverage: 11.7 in Q20.

* NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

* 1 13447: contig of 13447 bp in length
* 13448 13547: gap of 100 bp
* 13548 13569: contig of 2022 bp in length
* 13570 13569: gap of 100 bp
* 17887 17886: contig of 2217 bp in length
* 17887 17886: gap of 100 bp
* 20138 20137: contig of 2151 bp in length
* 20238 24761: contig of 4524 bp in length
* 24762 24861: gap of 100 bp
* 33737 33836: contig of 8875 bp in length
* 33837 40733: contig of 6897 bp in length
* 40734 40833: gap of 100 bp
* 40834 68844: contig of 28011 bp in length
* 68845 68944: gap of 100 bp
* 68945 79259: contig of 10315 bp in length
* 79260 79359: gap of 100 bp
* 79360 93275: contig of 13916 bp in length
* 93276 93375: gap of 100 bp
* 93376 100339: contig of 6963 bp in length
* 100340 100439: gap of 100 bp
* 100440 106496: contig of 6058 bp in length
* 106497 106596: gap of 100 bp
* 106597 114601: contig of 8005 bp in length
* 114602 114701: gap of 100 bp
* 135194 135193: contig of 20492 bp in length
* 135194 135293: gap of 100 bp
* 135294 153360: contig of 18067 bp in length
* 153361 153460: gap of 100 bp

```

```

* 153461 170470: contig of 17010 bp in length
* 170471 170570: gap of 100 bp
* 170571 174711: contig of 4141 bp in length.
FEATURES
    source
        1..174711
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="17"
            /map="17"
            /clone="RP11-634P5"
            /clone_lib="RPC1-11 Human Male BAC"
            /notes="assembly_fragment"
            /clone_end:SP6
            vector_side:left
            13448..13547
                /estimated_length=100
            13548..15569
                /notes="assembly_fragment"
            15570..15669
                /estimated_length=100
            15670..17886
                /notes="assembly_fragment"
            17887..17986
                /estimated_length=100
            17987..20137
                /notes="assembly_fragment"
            20138..20237
                /estimated_length=100
            20238..24761
                /notes="assembly_fragment"
            24762..24861
                /estimated_length=100
            24862..33736
                /notes="assembly_fragment"
            33737..33836
                /estimated_length=100
            33837..40733
                /notes="assembly_fragment"
            40734..40833
                /estimated_length=100
            40834..68844
                /notes="assembly_fragment"
            68845..68944
                /estimated_length=100
            68945..79259
                /notes="assembly_fragment"
            79260..79359
                /estimated_length=100
            79360..93275
                /notes="assembly_fragment"
            93276..93375
                /estimated_length=100
            93376..100339
                /notes="assembly_fragment"
            100340..100439
                /estimated_length=100
            100440..106496
                /notes="assembly_fragment"
            106497..106596
                /estimated_length=100
            106597..114601
                /notes="assembly_fragment"
            114602..114701
                /estimated_length=100
            114702..135193
                /notes="assembly_fragment"
            135194..135293
                /estimated_length=100
            135294..153360
                /note="assembly_fragment"

```

```
Query Match      24.9%; Score 92.2; DB 14; Length 174711;
Best Local Similarity 70.1%; Pred. No. 2.4e-18;
Matches 124; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 194 AGACTTCTTGATCTTCTCAGCATATGACTCTCATATGAACACTCTCTCTCGCTAG 253
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8124 AGGCTGTTCCTTGTCTTCTCATCTTACTCTCCACTACAAACACACCTCTCGCTG 8183

QY 254 GAGTGGCAGCTACATCGCTAATCCAGCAGCTTGGAGCTGAGCAGGAGGATTA 313
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8184 GTGTGGTGGCTGACATCTGTAATCCCGACATTTTGGGAGCTGAGCAGGAGGATCA 8243

QY 314 GAGCCAGGAGTTTGAGACAGGCTGGGCAATACAGCGACACTCTCTCTCTAAAA 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8244 GAGCCCAAGAGTTAGACACAGCTGGGCAACATAGCAGACATCTCTACAAA 8300

RESULT 9
AL355602      99163 bp      DNA      linear      PRI 18-MAY-2005
LOCUS
DEFINITION
Human DNA sequence from clone RP5-1166F10 on chromosome 1 Contains
the 3' end of a novel gene (FLJ36179), complete sequence.
ACCESSION
AL355602.11 GI:13560009
VERSION
HTG; FLJ36179.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 99163)
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Apr 6, 2001 this sequence version replaced gi:13396512.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RP5-1166F10 is from the library RPCI-5 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk

-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
Location/Qualifiers
1. 99163
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="RZPD:RPCI704F101166"
/db_xref="taxon:9606"
/chromosome="1"
/clones="RP5-1166F10"
/clone_lib="RPCI-5"

FEATURES
source
misc_feature
1
```

```
/note="Clone left end: RP5-1166F10"
join(AL591916.8:151507..152088,AL591916.8:155009..155043,
AL591916.8:155768..156004,5057..5484,5645..5745,
6291..6984)
/locus_tag="RP5-1166F10.1-001"
join(AL591916.8:151507..152088,AL591916.8:155009..155043,
AL591916.8:155768..156004,5057..5484,5645..5745,
6291..6984)
/locus_tag="RP5-1166F10.1-001"
/product="novel transcript"
/note="match: ESTs: Em:AW628519.1
match: cDNAs: Em:AK093498.1 Em:AX748168.1"
6956..6961
/locus_tag="RP5-1166F10.1-001"
6984
/locus_tag="RP5-1166F10.1-001"
99064
/note="Clone left end: RP1-37J18"

ORIGIN
Query Match      24.8%; Score 91.8; DB 8; Length 99163;
Best Local Similarity 77.6%; Pred. No. 3.1e-18;
Matches 111; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 228 TCATATGAAACTCTCTCTGCTAGGAGTGGCAGCTACATCGCTAATCCAGCACCTT 287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12266 TCTTATAAAGAGACCCCTGCTGGCATGTGGCTCACAATGTATCCAGCACCTTT 12325

QY 288 GGAAGCTGAGCAGGAGGATTACTGGAGCCAGGAGTTTGAGACAGCGCTGGCAATAC 347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12326 GGGAGCCGAGGAGGAGATCACTGGAGCCAGGAGTTTCAAGACAGCGTGGGTAACAC 12385

QY 348 AGGAGACTCTCTCTCTAAAAA 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12386 AGCAGACTCTGTCTCAAAAAA 12408

RESULT 10
AC024517      165079 bp      DNA      linear      HTG 21-MAR-2000
LOCUS
DEFINITION
Homo sapiens chromosome 1 clone RP11-168B8 map 1, WORKING DRAFT
SEQUENCE, 30 unordered pieces.
ACCESSION
AC024517.2 GI:7272182
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 165079)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepe,Y., Collangelo,M., Collins,S., Collamore,A., Cooke,P.,
DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M.,
Feneator,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
Galgani,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Landers,T., Laroque,K., Lehotzky,J., Levine,R.,
Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Olivari,T.M.,
Peterson,K., Pierre,N., Pisan,C., Pollara,V., Raymond,C.,
Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
```

Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.
 Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
 Zody,M.

TITLE Direct Submission
JOURNAL Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Mar 21, 2000 this sequence version replaced gi:7108313.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L7252
 Center clone name: 168.B.8
 ----- Summary Statistics
 Sequencing vector: M13; M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 144165 bases at least Q40
 Consensus quality: 153730 bases at least Q30
 Consensus quality: 158511 bases at least Q20
 Insert size: 170000; agarose-fp
 Insert size: 162179; sum-of-contigs
 Quality coverage: 3.2 in Q20 bases; agarose-fp
 Quality coverage: 3.4 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 30 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

*	1	1125:	contig of 1125 bp in length
*	1126	1225:	gap of 100 bp
*	1226	2573:	contig of 1348 bp in length
*	2574	2673:	gap of 100 bp
*	2674	3825:	contig of 1152 bp in length
*	3826	3925:	gap of 100 bp
*	3926	5196:	contig of 1271 bp in length
*	5197	5296:	gap of 100 bp
*	5297	6379:	contig of 1083 bp in length
*	6380	6479:	gap of 100 bp
*	6480	8499:	contig of 2020 bp in length
*	8500	8599:	gap of 100 bp
*	8600	11025:	contig of 2426 bp in length
*	11026	11125:	gap of 100 bp
*	11126	13438:	contig of 2313 bp in length
*	13439	13538:	gap of 100 bp
*	13539	16281:	contig of 2743 bp in length
*	16282	16381:	gap of 100 bp
*	16382	18437:	contig of 2056 bp in length
*	18438	18537:	gap of 100 bp
*	18538	22277:	contig of 3740 bp in length
*	22278	22377:	gap of 100 bp
*	22378	25101:	contig of 2724 bp in length
*	25102	25201:	gap of 100 bp
*	25202	27806:	contig of 2605 bp in length
*	27807	27906:	gap of 100 bp
*	27907	30832:	contig of 2926 bp in length
*	30833	30933:	gap of 100 bp
*	30933	33714:	contig of 2782 bp in length
*	33715	33814:	gap of 100 bp
*	33815	36617:	contig of 2803 bp in length
*	36618	36717:	gap of 100 bp
*	36718	39412:	contig of 2695 bp in length
*	39413	39512:	gap of 100 bp
*	39513	43281:	contig of 3769 bp in length

```

gap          18438..18537
            /estimated_length=100
misc_feature 18538..22277
            /note="assembly_fragment"
gap          22278..22377
            /estimated_length=100
misc_feature 22378..25101
            /note="assembly_fragment"
gap          25102..25201
            /estimated_length=100
misc_feature 25202..27806
            /note="assembly_fragment"
gap          27807..27906
            /estimated_length=100
misc_feature 27907..30832
            /note="assembly_fragment"
gap          30833..30932
            /estimated_length=100
misc_feature 30933..33714
            /note="assembly_fragment"

Query Match      24.8%; Score 91.8; DB 14; Length 165079;
Best Local Similarity 77.6%; Pred. No. 3.3e-18;
Matches 111; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 228 TCATATGAACTCTCTCTGGTAGGAGTGCGAGCTACATCGCTAATCCAGCACCTT 287
Db 45608 TCTTATAAAGAGACCCCTGGCTGGCATGGTGCATCTGTAATCCAGCACCTT 45667

QY 288 GGGAAGCTGAGGAGGAGGATTACTGGAGGCCAGGAGTTTGAGACAGCGTGGGCAATAC 347
Db 45668 GGGAGGCCGAGGAGGAGATCTAGGAGGCCAGGAGTTCAAGACCGTGGGTAAAC 45727

QY 348 AGCGAGACTCTCTCTCTCTAAAAA 370
Db 45728 AGCAAGACTGTCTCAAAAAA 45750

RESULT 11
AC021394/c
LOCUS      167300 bp      DNA      linear      HTG 10-MAR-2001
DEFINITION Homo sapiens clone RP11-28E15, WORKING DRAFT SEQUENCE, 11 unordered
            pieces.
ACCESSION  AC021394
VERSION     AC021394.3  GI:7630674
KEYWORDS   HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE  1 (bases 1 to 167300)
AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE     Homo sapiens, clone RP11-28E15
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 167300)
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguelavkiy,L., Bouhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Donino,M., Doyle,M., Feneator,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galegan,J.,
            Gardyna,S., Grant,G., Hagos,B., Hesford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.

```

TITLE JOURNAL COMMENT

Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 21, 2000 this sequence version replaced gi:6850427.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence submissions@genome.wi.mit.edu

----- Project Information

Center project name: L4746

Center clone name: 28_E_15

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 158720 bases at least Q40

Consensus quality: 162531 bases at least Q30

Insert size: 167000; agarose-fp

Insert size: 166300; sum-of-contigs

Quality coverage: 4.7 in Q20 bases; agarose-fp

Quality coverage: 4.7 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

* 1 1018: contig of 1018 bp in length
* 1019 1118: gap of 100 bp
* 1119 4175: contig of 3057 bp in length
* 4176 4275: gap of 100 bp
* 4276 8989: contig of 4714 bp in length
* 8990 9089: gap of 100 bp
* 9090 14343: contig of 5254 bp in length
* 14344 14443: gap of 100 bp
* 14444 20389: contig of 5946 bp in length
* 20390 20489: gap of 100 bp
* 20490 28508: contig of 8019 bp in length
* 28509 28608: gap of 100 bp
* 28609 37887: contig of 9279 bp in length
* 37888 37988: gap of 100 bp
* 37989 54331: contig of 16344 bp in length
* 54332 54431: gap of 100 bp
* 54432 84316: contig of 29885 bp in length
* 84317 121379: contig of 36963 bp in length
* 121380 121479: gap of 100 bp
* 121480 167300: contig of 45821 bp in length.

FEATURES

source

1..167300
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-28E15"

/clone_lib="RPC1-11 Human Male BAC"
1..1018

misc_feature

1019..1118
/note="assembly_fragment"

gap

/estimated_length=100

misc_feature

1119..4175
/note="assembly_fragment"

gap

4176..4275
/estimated_length=100

misc_feature

4276..4898
/note="assembly_fragment"


```

gap      8990..9089
misc_feature /estimated_length=100
9090..14343
/note="assembly_fragment"
gap      14344..14443
misc_feature /estimated_length=100
14444..20389
/note="assembly_fragment"
clone_end:T7
vector_side:left"
gap      20390..20489
misc_feature /estimated_length=100
20490..28508
/note="assembly_fragment"
gap      28509..28608
misc_feature /estimated_length=100
28609..37887
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
gap      37888..37987
misc_feature /estimated_length=100
37988..54331
/note="assembly_fragment"
gap      54332..54431
misc_feature /estimated_length=100
54432..84316
/note="assembly_fragment"
84317..84416
gap      /estimated_length=100
84417..121379
misc_feature /note="assembly_fragment"
121380..121479
gap      /estimated_length=100
121480..167300
misc_feature /note="assembly_fragment"

ORIGIN
Query Match      24.8%; Score 91.8; DB 14; Length 167300;
Best Local Similarity 77.6%; Pred No. 3.3e-18;
Matches 111; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 228 TCATATGAACCTCTCTCTGCTAGGAGTGGCAGCCTACATCGCTAATCCAGCACCTT 287
Db 13204 TCTTATAAAGAGACCCCTGGCTGGCATGCTGCATCTGTATATCCAGCACCTT 13145

Qy 288 GGGAGCTGAGCAGGAGATTACTGGAGGCCAGGAGTTTGAGACCAGGCTGGGCAATAC 347
Db 13144 GGGAGCGCAGGAGGAGATCACTGGAGCCAGGAGTTCAAGACCAGCGTGGTAAACAC 13085

Qy 348 AGCGAGACTCTCTCTCTAAATA 370
Db 13084 AGCAAGACTCTGTCTCAAAAAA 13062

RESULT 12
AC112211/c AC112211 182509 bp DNA linear PRI 24-MAY-2002
LOCUS Homo sapiens chromosome 3 clone RP11-627J17, complete sequence.
DEFINITION AC112211
ACCESSION AC112211
VERSION AC112211.2 GI:21166208
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 182509)
Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Raymond,C. and
Haugen,E.D.
Direct Submission
REFERENCE 1 (bases 1 to 182509)
AUTHORS Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
TITLE Submitted (24-MAY-2002) Genome Center, University of Washington,
JOURNAL Box 352145, Seattle, WA 98195, USA
2 (bases 1 to 182509)
Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Raymond,C. and
Haugen,E.D.
Direct Submission
TITLE Submitted (24-MAY-2002) Genome Center, University of Washington,
JOURNAL Box 352145, Seattle, WA 98195, USA
COMMENT On May 24, 2002 this sequence version replaced gi:18767508.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchgs@u.washington.edu
----- Project Information
Center project name: chr-3
Center clone name: RP11-627J17 (bc0502)
----- Summary Statistics
Sequencing vector: plasmid; 100% of reads
Chemistry: Dye-terminator ET; 50% of reads
Chemistry: Dye-terminator Big Dye; 50% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 182361 bases at least Q40
Consensus quality: 182497 bases at least Q30
Consensus quality: 182509 bases at least Q20
Insert size: 182509; sum-of-contigs
Quality coverage: 7.9x in Q20 bases; sum-of-contigs
-----
Overlapping Sequences:
5': Mapping in progress
3': Mapping in progress
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.
-----
Sequence Validation:
This sequence has been validated by Multiple Complete Digest
fingerprinting. Comparison of the experimentally derived digest
fragments with sequence-predicted fragments is given below.
The electronically-digested sequence consists of both insert and
vector, in order to accurately represent the entire circular BAC.
Small fragments below a variable cutoff (approximately 400-800 bp)
are not resolved in the fingerprint and hence do not appear
in the table. There are no significant remaining discrepancies
between the experimental and predicted values. Uniquely ordered
fragments are separated by dashed lines.
ECORI
SeqDerMap FngPrnt SeqDerMap FngPrnt SeqDerMap FngPrnt
-----
8696 8777 4710 4671 2282 2251
-----

```

```

REFERENCE 2 (bases 1 to 182509)
AUTHORS Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 182509)
Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Raymond,C. and
Haugen,E.D.
Direct Submission
TITLE Submitted (24-MAY-2002) Genome Center, University of Washington,
JOURNAL Box 352145, Seattle, WA 98195, USA
COMMENT On May 24, 2002 this sequence version replaced gi:18767508.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchgs@u.washington.edu
----- Project Information
Center project name: chr-3
Center clone name: RP11-627J17 (bc0502)
----- Summary Statistics
Sequencing vector: plasmid; 100% of reads
Chemistry: Dye-terminator ET; 50% of reads
Chemistry: Dye-terminator Big Dye; 50% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 182361 bases at least Q40
Consensus quality: 182497 bases at least Q30
Consensus quality: 182509 bases at least Q20
Insert size: 182509; sum-of-contigs
Quality coverage: 7.9x in Q20 bases; sum-of-contigs
-----
Overlapping Sequences:
5': Mapping in progress
3': Mapping in progress
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.
-----
Sequence Validation:
This sequence has been validated by Multiple Complete Digest
fingerprinting. Comparison of the experimentally derived digest
fragments with sequence-predicted fragments is given below.
The electronically-digested sequence consists of both insert and
vector, in order to accurately represent the entire circular BAC.
Small fragments below a variable cutoff (approximately 400-800 bp)
are not resolved in the fingerprint and hence do not appear
in the table. There are no significant remaining discrepancies
between the experimental and predicted values. Uniquely ordered
fragments are separated by dashed lines.
ECORI
SeqDerMap FngPrnt SeqDerMap FngPrnt SeqDerMap FngPrnt
-----
8696 8777 4710 4671 2282 2251
-----

```


REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

2 (bases 1 to 101769)
DOE Joint Genome Institute.
Direct Submission
Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

3 (bases 1 to 101769)
DOE Joint Genome Institute.
Direct Submission
Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Mar 7, 2002 this sequence version replaced gi:15426059.

-----Genome Center
Center: Joint Genome Institute
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 285795
Center clone name: CIT-HSPC_262A8

Summary Statistics
Consensus quality: 101567 bases at least Q40
Consensus quality: 101611 bases at least Q30
Consensus quality: 101641 bases at least Q20
Estimated insert size: 104000; agarose-fp estimation
Estimated insert size: 101669; sum-of-contigs estimation
Quality coverage: 34.54 in Q20 bases; agarose-fp estimation
Quality coverage: 35.33 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 12059: contig of 12059 bp in length
* 12060 12159: gap of unknown length
* 12160 101769: contig of 89610 bp in length.

Location/Qualifiers
1. 101769
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTC-262A8"
/clone_lib="CalTech human BAC library C"
12060. .12159
/estimated_length=unknown

gap

ORIGIN

Query Match 24.6%; Score 91.2; DB 14; Length 101769;
Best Local Similarity 71.4%; Pred. NO. 4.9e-18;
Matches 137; Conservative 0; Mismatches 48; Indels 7; Gaps 1;

QY 2 GGTTCAGCCACGCAAGTGTAGAGTCTACAC-----CCTCTATACAGGATACA 54
DB 74874 GGTTCAGCCACGCAAGTGTAAAGAGCTACAAATATCTCCAAATATGAAATCAC 74815

QY 55 GCAATGTGAAACACAGAGTCTATGCACTGAGGTTTCATCAAAACACTGGAAGCCATGGTA 114
DB 74814 AGACACTGGTCAGAGGCCAAGCACTGAGGTTCTGTCAAGCTCTCGAAGCCCTGGTA 74755

QY 115 GTTATGTCCTTGTAGTAGTATGTCAACTCTCAACCAATGTCCTTCCTGGCTGATGGAGA 174
DB 74754 GTTACTGTCTTAATGTGTATGTACAGTCTCAACCAATGTCCTTCCTGGCTGATGGAGA 74695

QY 175 AATGGCTGCC 186
DB 74694 AAGGATGCC 74683

RESULT 14
AC025473/c
LOCUS
DEFINITION

AC025473
VERSION
KEYWORDS
SOURCE
ORGANISM

AC025473.2 GI:7417674
HTG; HTGS_PHASE1; HTGS_DRAFT.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Suarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 128540)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 5
Unpublished
2 (bases 1 to 128540)
DOE Joint Genome Institute.
Direct Submission
Submitted (09-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Apr 5, 2000 this sequence version replaced gi:7211967.

-----Genome Center
Center: Joint Genome Institute
Web site: <http://www.jgi.doe.gov>

-----Summary Statistics
Consensus quality: 118729 bases at least Q40
Consensus quality: 125307 bases at least Q30
Consensus quality: 126460 bases at least Q20
Estimated insert size: 128540; sum-of-contigs estimation
Estimated insert size: 202600; agarose-fp estimation
Quality coverage: 3.58x in Q20 bases; agarose-fp estimation
Quality coverage: 5.64x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 2361: contig of 2361 bp in length
* gap of unknown length
* 2362 4766: contig of 2405 bp in length
* gap of unknown length
* 4767 7300: contig of 2534 bp in length
* gap of unknown length
* 7301 10400: contig of 3100 bp in length
* gap of unknown length
* 10401 14252: contig of 3852 bp in length
* gap of unknown length
* 14253 20943: contig of 6691 bp in length
* gap of unknown length
* 20944 26181: contig of 5238 bp in length
* gap of unknown length
* 26182 32137: contig of 5956 bp in length
* gap of unknown length
* 32138 47316: contig of 15179 bp in length
* gap of unknown length
* 47317 71055: contig of 23739 bp in length
* gap of unknown length
* 71056 128540: contig of 57485 bp in length.

Location/Qualifiers
1. 128540
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-183N2"

FEATURES
source

ORIGIN

Query Match 24.6%; Score 91.2; DB 14; Length 128540;
Best Local Similarity 71.4%; Pred. No. 5e-18; Indels 7; Gaps 1;
Matches 137; Conservative 0; Mismatches 48; Indels 7; Gaps 1;

QY 2 GGTTCAGCCACAGCAAAATGTATAGAGTGCTACAAC-----CCTCTATACAGGGATACA 54
DB 128064 GGTTCAGCCACAGCAAAATGTATAGAGTGCTACAAC-----CCTCTATACAGGGATACA 54

QY 55 GCAATGTGAACACGAGGTCATGCTGAGGTTTTCATCAAAACACTGGAAGCCCATGGTA 114
DB 128004 GGTTCAGCCACAGCAAAATGTATAGAGTGCTACAAC-----CCTCTATACAGGGATACA 114

QY 115 GTTATTGTCCTTGATAGGTCATGCTGAGGTTTTCATCAAAACACTGGAAGCCCATGGTA 174
DB 127944 GTTATTGTCCTTGATAGGTCATGCTGAGGTTTTCATCAAAACACTGGAAGCCCATGGTA 174

QY 175 AATGGGCTGCCC 186
DB 127884 AAGGGATGGCCC 127873

RESULT 15
AC073542 141633 bp DNA linear PRI 01-MAY-2001
LOCUS Homo sapiens chromosome 5 clone RP11-221L16, complete sequence.
DEFINITION AC073542
ACCESSION AC073542
VERSION AC073542.4 GI:13876495
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 141633)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 141633)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (22-JUN-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 3 (bases 1 to 141633)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (28-APR-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 4 (bases 1 to 141633)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT On Apr 28, 2001 this sequence version replaced gi:13677059.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.sngc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.1.

FEATURES
source
1. .141633
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-221L16"

ORIGIN

Query Match 24.6%; Score 91.2; DB 8; Length 141633;
Best Local Similarity 71.4%; Pred. No. 5e-18; Indels 7; Gaps 1;
Matches 137; Conservative 0; Mismatches 48; Indels 7; Gaps 1;

QY 2 GGTTCAGCCACAGCAAAATGTATAGAGTGCTACAAC-----CCTCTATACAGGGATACA 54
DB 66221 GGTTCAGCCACAGCAAAATGTATAGAGTGCTACAAC-----CCTCTATACAGGGATACA 54

QY 55 GCAATGTGAACACGAGGTCATGCTGAGGTTTTCATCAAAACACTGGAAGCCCATGGTA 114
DB 66281 AGACACTTGGTCAGAGGCCCAAGCAACTGAGGTTTCTGTCAAAGCTCTGGAAGCCCATGGTA 114

QY 115 GTTATTGTCCTTGATAGGTCATGCTGAGGTTTTCATCAAAACACTGGAAGCCCATGGTA 174
DB 66341 GTTATTGTCCTTGATAGGTCATGCTGAGGTTTTCATCAAAACACTGGAAGCCCATGGTA 174

QY 175 AATGGGCTGCCC 186
DB 66401 AAGGGATGGCCC 66412

Search completed: December 7, 2005, 04:21:46
Job time : 3280.23 secs


```
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239933P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241788P.
PR 20-OCT-2000; 2000US-0241789P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system cancers
XX and metastases.
XX Disclosure; SEQ ID NO 11655; 1701pp + Sequence Listing; English.
XX The invention relates to novel genes (AB11004-ABA21534) and proteins
XX (AB14678-AB18001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemias; (d) wound healing
XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 27733 BP; 7380 A; 6291 C; 6413 G; 7649 T; 0 U; 0 Other;

Query Match 25.2%; Score 93.4; DB 5; Length 27733;
Best Local Similarity 80.7%; Pred. No. 2.6e-20;
Matches 109; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 50 ATACAGCAATGTGAAACGAGAGTCATGTCATCTGAGGTTTCATCAAAACACTGGAAGCCA 109
DB 4609 ATAGTAGACTGTGGCCAGAGGCCATGTCATCTGAGGTTTGTCCAAACACTGGAAGCCC 4668
QY 110 TGGTAGTTATTGTCCTTGATAGGTATGTCAACCTCAACACCAATGTCCTTCTGCTGAT 169
DB 4669 TGGTAGTTATTGTCCTTCTTTAGGTATGTCAACCTCAACACCAATGTCCTTCTGCTGAT 4728
QY 170 GGAGAAATGGGCTGC 184
DB 4729 GGAGAAACAGACAGC 4743

RESULT 3
ACF87264/c
ID ACF87264 standard; DNA; 456 BP.
XX ACF87264;
AC ACF87264;
XX 02-JUN-2005 (first entry)
XX Human SIRS/sepsis diagnostic marker DNA fragment 6124.
XX Systemic inflammatory response syndrome; SIRS; antibacterial;
XX immunosuppressive; antiinflammatory; diagnosis; sepsis; ds.
XX Homo sapiens.
XX WO2004087949-A2.
XX 14-OCT-2004.
XX 31-MAR-2004; 2004WO-EP003419.
XX 02-APR-2003; 2003DE-01015031.
XX 08-AUG-2003; 2003DE-01036511.
XX 02-SEP-2003; 2003DE-01040395.
XX
```

PA (SIRS-) SIRS LAB GMBH.
XX Russwurm S, Reinhart K, Saluz H, Straube E, Zipfel PF, Deigner H;
XX WPI; 2004-748070/73.
DR In vitro detection of systemic inflammatory response syndrome and related
XX conditions, for e.g. monitoring progression, comprises detecting abnormal
XX expression of disease-related genes.
PT Disclosure; Page; 75pp; German.
XX
XX The invention relates to a novel method for in vitro detection of
XX systemic inflammatory response syndrome (SIRS). The method comprises
XX detecting abnormal expression of disease-related genes, or their
XX associated peptides. The method of the invention demonstrates
XX antibacterial, immunosuppressive and antiinflammatory applications and
XX may be used for early differential diagnosis, monitoring progression, and
XX assessing risk, assessing the likely response to treatment and for post
XX mortem diagnosis of systemic inflammatory response syndrome, sepsis and
XX sepsis-like conditions. The recombinant or synthetic nucleic acid
XX sequences of the invention, or derived proteins or peptides, may be
XX useful as calibrants in assays for the specified diseases, for evaluating
XX activity or toxicity in screening for active agents and/or for
XX preparation of agents for treatment or prevention of the specified
XX diseases. The current sequence is that of a human SIRS/sepsis diagnostic
XX marker DNA fragment of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at ftp.wipo.int/pub/published
XX pct sequences. Furthermore, a number of arbitrary SEQ ID NO.s are
XX disclosed within the specification, however, these have not been taken
XX into account during indexing due to inconsistencies in application and
XX format
XX
SQ Sequence 456 BP; 107 A; 87 C; 134 G; 123 T; 0 U; 5 Other;

Query Match 24.2%; Score 89.6; DB 13; Length 456;
Best Local Similarity 67.2%; Pred. No. 7e-20; Indels 0; Gaps 0;
Matches 125; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 31 TACAACCTCTTATACAGGGATACAGCAATGTGAACCCAGAGTCATCTCAGGTTTC 90
DB 236 TGCTCCCAACAATATAAATGTCAGANTGTTGCCAGAGGCCACATGTCAGGTTTC 177

QY 91 ATCAAAACACTGGAGCCATGGTAGTTATTGTCCTGTAGTAGTATGTCAACCTCAACACC 150
DB 176 ATGAAGACACTGGAAACCATGGTAGTTACTGTTCTGTAAAGTATGTGGCCTCAACACC 117

QY 151 AATGTCCTTCTGCTGATGGAGAAATGGGCTGCCCGATTTCACAGACTTCTTGATCTTC 210
DB 116 AGTGCCTTCTGCTGATGAGAAATGGATGGCACTTGCCCTGAAAGGATGTCAGCTG 57

QY 211 CTCAGC 216
DB 56 CCCAGC 51

RESULT 4
ADH10017_1/c
Continuation (2 of 4) of ADH10017 from base 100001 (Human chromosome 2p21-22 fragment cd
WP Sequence split into 4 fragments LOCUS ADH10017 Accession Adh10017
WP Fragment Name Begin End
WP ADH10017_0 1 110000
WP ADH10017_1 100001 210000
WP ADH10017_2 200001 310000
WP ADH10017_3 300001 365186

Query Match 23.8%; Score 88; DB 10; Length 110000;
Best Local Similarity 80.5%; Pred. No. 3.7e-16; Indels 0; Gaps 0;
Matches 103; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 243 CTCCTGGCTAGGAGTGGCAGCTACATCGGTATATCCAGCACCTTTGGGAAGCTGAGGCAG 302
DB 105042 CACCTGGCTAGGAGTGGCAGCTACATCGGTATATCCAGCACCTTTGGGAAGCTGAGGCAG 302

DB 24799 CACCTGGCTGGGCATGGTGGCTCACATCTATAATCCAGCACCTTTGGGAGACCGAGGCAG 24740
QY 303 GAGGATTACTGGAGCCAGGAGTTTGAGACCAGGCTGGGCAATACAGCGAGACTCTCTCT 362
DB 24739 GAGGATCACTTGAGCCAGGACTTTTGAGACCAGGCTTTGAGAGACTCTCTCTCTCT 24680
QY 363 CCTAAAAA 370
DB 24679 CTACAAAA 24672

RESULT 5
ADQ97266_0/c
WP Sequence split into 4 fragments LOCUS ADQ97266 Accession Adq97266
WP Fragment Name Begin End
WP ADQ97266_0 1 110000
WP ADQ97266_1 100001 210000
WP ADQ97266_2 200001 310000
WP ADQ97266_3 300001 355211
ID ADQ97266 standard; DNA; 355211 BP.
XX
XX ADQ97266;
XX
XX 07-OCT-2004 (first entry)
XX Human cancer associated sequence HD08-023, SEQ ID 242.
XX
XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
XX
XX Homo sapiens.
XX
XX WO2004060304-A2.
XX
XX 22-JUL-2004.
XX
XX 22-DEC-2003; 2003WO-US041389.
XX
XX 27-DEC-2002; 2002US-00330773.
XX
XX (SAGR-) SAGRES DISCOVERY INC.
XX
XX Morris DW, Malandro MS;
XX WPI; 2004-543781/52.
XX
XX New isolated cancer associated nucleic acids comprising at least 10
XX contiguous nucleotides, useful for diagnosing, preventing and/or treating
XX cancers such as leukemia and lymphoma.
XX
XX Claim 1; SEQ ID NO 242; 199pp; English.
XX
XX The present invention relates to cancer associated sequences (ADQ97025-
XX ADQ98004). The sequences are useful for the diagnosis, prevention and/or
XX treatment of cancer, such as leukemia and lymphoma. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 355211 BP; 96742A; 68188C; 72417G; 106131T; 0U; 117330Other;
XX
Query Match 23.8%; Score 88; DB 12; Length 110000;
Best Local Similarity 80.5%; Pred. No. 3.7e-18; Indels 0; Gaps 0;
Matches 103; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 243 CTCCTGGCTAGGAGTGGCAGCTTACATCGCTATATCCAGCACCTTTGGGAAGCTGAGGCAG 302
DB 105102 CACCTGGCTGGGCATGGTGGCTCACATCTATAATCCAGCACCTTTGGGAGACCGAGGCAG 105043

QY 303 GAGGATTACTGGAGCCAGGAGTTTGAGACCAGGCTGGGCAATACAGCGAGACTCTCTCT 362
DB 105042 GAGGATCACTTGAGCCAGGACTTTTGAGACCAGGCTTTGAGAGACTCTCTCTCTCT 104983
QY 363 CCTAAAAA 370

PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232081P.
PR 08-SEP-2000; 2000US-0232082P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.

PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-502630/55.
XX Polynucleotides encoding digestive system antigens, useful for
diagnosing, treating, preventing and/or prognosing disorders of the
digestive system, particularly cancer and cancer metastases.
XX Disclosure; SEQ ID NO 3133; 986pp; English.
XX The present invention provides the protein and coding sequences of a
number of human digestive system antigens. These can be used in the
diagnosis, treatment and prevention of digestive system disorders,
including cancer, Meckel's diverticulum, bacterial or parasitic
infections, appendicitis, Hirschsprung's disease, chronic colitis or
ulcerative colitis. The present sequence is a genomic DNA fragment
encoding a digestive system antigen of the invention
XX
SQ Sequence 13403 BP; 3792 A; 3003 C; 2921 G; 3687 T; 0 U; 0 Other;
Query Match 23.5%; Score 86.8; DB 4; Length 13403;
Best Local Similarity 76.8%; Pred. No. 3.4e-18;
Matches 106; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
Qy 233 TGAACCTCTCTCTGGCTAGGAGTGGCAGCTACATCGCTTAATCCAGCACCTTGGAA 292
Db 7263 TAAATTTCTCACTTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 7204
Qy 293 GCTGAGGCAGGAGATTACTGGAGGCCAGGAGTTTGAGACACGCTGGGCAATACAGCA 352
Db 7203 GCTGAGGCAAGTGGATCACTAGAGTCCAGGAGTTAGAGACCAGCTGGCAGCATAGTGA 7144
Qy 353 GACTCTCTCTCTTAAAAA 370
Db 7143 GACCTGTCTCTACAAAA 7126
RESULT 9
ABK42627/c
ID ABK42627 standard; DNA; 23378 BP.
XX
XX AC ABK42627;
XX
XX DT 21-MAY-2002 (first entry)
XX
XX DE Genomic sequence #526 encoding novel human connective tissue polypeptide.
XX
XX KW Human; connective tissue related disorder; cancer; gene therapy;

KW cytostatic; gene; ds.
XX Homo sapiens.
OS
XX
PN WO200153343-A1.
XX
PD 02-AUG-2001.
XX
XX
PF 17-JAN-2001; 2001WO-US001322.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0248525P.
PR 08-NOV-2000; 2000US-0248526P.
PR 08-NOV-2000; 2000US-0248527P.
PR 08-NOV-2000; 2000US-0248528P.
PR 08-NOV-2000; 2000US-0248532P.
PR 08-NOV-2000; 2000US-0248609P.
PR 08-NOV-2000; 2000US-0248610P.
PR 08-NOV-2000; 2000US-0248611P.
PR 08-NOV-2000; 2000US-0248613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

```
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-565190/63.
XX DR
XX PT Nucleic acid encoding novel connective tissue associated polypeptides,
XX PT used in diagnosing, preventing, treating or ameliorating a disorder such
XX PT as cancer or rheumatoid arthritis.
XX PS Disclosure; SEQ ID NO 1514; 673pp; English.
XX CC
XX CC The present invention relates to the isolation of novel human connective
XX CC tissue related polypeptides (AAU86435-AAU86923) and the polynucleotide
XX CC (cDNA and genomic) sequences encoding them. The sequences of the
XX CC invention are useful in the diagnosis, treatment, prevention and/or
XX CC prognosis of diseases associated with connective tissue(s), including
XX CC cancer. The polynucleotide sequences of the invention are also useful in
XX CC gene therapy. ABK42102-ABK43116 represent genomic sequences encoding the
XX CC novel human connective tissue related polypeptides. Note: The sequence
XX CC data for this patent did not form part of the printed specification, but
XX CC was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ
SQ Sequence 23378 BP; 6470 A; 5283 C; 4895 G; 6730 T; 0 U; 0 Other;

Query March 23.5%; Score 86.8; DB 4; Length 23378;
Best Local Similarity 76.8%; Pred. No. 4.4e-18;
Matches 106; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 233 TGAACCTCTCTCTGGCTAGGAGTGGCAGCTACATCGCTAATCCAGCACCTTGGGAA 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17240 TAAATTTCTTCACTTGGCTGGTGTGGTCTCAGCCTGTATATCCAGCACTTGGAG 17181

QY 293 GCTGAGGCAGGAGGATTACTGGAGCCAGGAGTTTGAGACCAAGCTGGGCAATACAGCA 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17180 GCTGAGGCAAGTGGATCACTAGAGTCCAGGAGTTAGAGACCAAGCTGGGCGCATAGTGA 17121

QY 353 GACTCTCTCTCTAAAAA 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 17120 GACCTGTCTTACAAAA 17103

RESULT 10
AAK89555/c
ID AAK89555 standard; DNA; 23378 BP.
XX AC
XX AC AAK89555;
XX DT
XX DT 05-NOV-2001 (first entry)
XX DE
XX DE Human digestive system antigen genomic sequence SEQ ID NO: 3131.
XX KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
XX KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
XX KW digestive system disorder; Meckel's diverticulum; ds.
XX OS Homo sapiens.
XX PN WO200155314-A2.
XX XX
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001324.
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-0190076P.
XX PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216800P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
```

PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-02411785P.
PR 20-OCT-2000; 2000US-02411786P.
PR 20-OCT-2000; 2000US-02411787P.
PR 20-OCT-2000; 2000US-02411808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249237P.
PR 17-NOV-2000; 2000US-0249239P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250150P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI, 2001-502630/55.
XX
PT Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX
PS Disclosure; SEQ ID NO 3131; 986pp; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic

CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a genomic DNA fragment
CC encoding a digestive system antigen of the invention
XX
SQ Sequence 23378 BP; 6470 A; 5283 C; 4895 G; 6730 T; 0 U; 0 Other;
Query Match 23.5%; Score 86.8; DB 4; Length 23378;
Best Local Similarity 76.8%; Pred. No. 4.4e-18;
Matches 106; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 233 TGAACCTCTCCTCTGGCTAGGAGTGGCAGCCTACATCCTAATCCAGACCTTGGGAA 292
Db 17240 TAAATTTCTCCACTTGGCTGGTGGTGGCTCAGCCTGTAAATCCAGACCTTTGGGAG 17181
QY 293 GCTGAGGCGAGGAGATTACTGGAGGCCAGGAGTTTGAGACCAGCTGGGGCNAACAGCGA 352
Db 17180 GCTGAGGCGAAGTGGATCACTAGAGTCCAGGAGTTAGAGACCGCTGGGCGCATAGTGA 17121
QY 353 GACTCTCTCTCTCTTAAAAA 370
Db 17120 GACCTGTCTCTTACAAAA 17103
RESULT 11
ADB60783/c
ID ADB60783 standard; DNA; 23378 BP.
XX ADB60783;
XX
DT 04-DEC-2003 (first entry)
XX
DE Connective tissue related genomic DNA #526.
XX
KW cytostatic; neuroprotective; nootropic; antiparkinsonian; cardiovascular;
KW antiarteriosclerotic; immunosuppressive; antirheumatic; antiarthritic;
KW antiinflammatory; antiallergic; antiasthmatic; dermatological;
KW nephrotic; virucide; fungicide; antibacterial; antiparasitic;
KW gene therapy; ds; connective tissues disorder; rheumatoid arthritis;
KW systemic lupus erythematosus; scleroderma; Sjogren's syndrome; cancer;
KW cancer metastasis; neoplasmia; leukemia; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; cardiovascular disease;
KW atherosclerosis; myocarditis; cardiopulmonary bypass complication;
KW autoimmune disease; multiple sclerosis; allergic reaction; asthma;
KW rhinitis; eczema; inflammatory condition; Crohn's disease; nephritis;
KW gastrointestinal disorder; inflammatory bowel disease;
KW organ transplant rejection; immune system disorder; Bruton's disease;
KW X-linked lymphoproliferative syndrome;
KW B-cell lymphoproliferative disorder; HIV; AIDS; infection;
KW chromosome identification; chromosome mapping;
KW connective tissue related polynucleotide; gene; ds.
XX Homo sapiens.
OS
XX
XX
XX US2003054375-A1.
XX
XX 20-MAR-2003.
XX
XX 07-MAR-2002; 2002US-00092154.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184564P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.

PR	11-JUL-2000;	2000US-0217496P.	PR	01-NOV-2000;	2000US-0244617P.
PR	14-JUL-2000;	2000US-0218290P.	PR	08-NOV-2000;	2000US-0246474P.
PR	26-JUL-2000;	2000US-0220963P.	PR	08-NOV-2000;	2000US-0246475P.
PR	26-JUL-2000;	2000US-0220964P.	PR	08-NOV-2000;	2000US-0246476P.
PR	14-AUG-2000;	2000US-0224518P.	PR	08-NOV-2000;	2000US-0246477P.
PR	14-AUG-2000;	2000US-0224519P.	PR	08-NOV-2000;	2000US-0246478P.
PR	14-AUG-2000;	2000US-0225213P.	PR	08-NOV-2000;	2000US-0246523P.
PR	14-AUG-2000;	2000US-0225214P.	PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000;	2000US-0225266P.	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000;	2000US-0225267P.	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000;	2000US-0225268P.	PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000;	2000US-0225270P.	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246611P.
PR	18-AUG-2000;	2000US-0226279P.	PR	08-NOV-2000;	2000US-0246613P.
PR	22-AUG-2000;	2000US-0226581P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226688P.	PR	17-NOV-2000;	2000US-0249208P.
PR	23-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249210P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249218P.
PR	08-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249246P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249300P.
PR	08-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250160P.
PR	12-SEP-2000;	2000US-0231968P.	PR	01-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-0232397P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232399P.	PR	05-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232400P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251869P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251990P.
PR	21-SEP-2000;	2000US-0234274P.	PR	11-DEC-2000;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234997P.	PR	05-JAN-2001;	2001US-0259678P.
PR	25-SEP-2000;	2000US-0234998P.			

CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
SQ Sequence 152330 BP; 40094 A; 34646 C; 37146 G; 40111 T; 0 U; 333 Other;
Query Match 23.1%; Score 85.6; DB 11; Length 152330;
Best Local Similarity 73.6%; Pred. No. 2.9e-17;
Matches 109; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
Qy 223 TACTCTCATATGAACACTCTCTCTGGCTAGGAGTGGCAGCCTACATCGCTAATCCAGC 282
Db 119498 TAGTACCCTTTATAAGAGGCGCCGAGGGCTGGCTGGTGCCTCACACCTATAATCCAGC 119439
Qy 283 ACCTTGGGAAGCTGAGGCAGCAGAGATTACTGGAGGCCAGGAGTTTGAGACCAGGCTGGGC 342
Db 119438 ACTTGGGAGGCTGAGGCAGCAGGAGTTGCTGGAGCCCGAGGAGTTGGAGACCAGCCTGGGC 119379
Qy 343 AATACAGCGAGACTCTCTCTCTAAATAA 370
Db 119378 AGCACAGTGAGACCCCATCTCCACAAAA 119351

Search completed: December 7, 2005, 00:02:54
Job time : 585.669 secs

Db
451 ATAATTCTAGCACTTTGGGAGGCTCAGGCAGGAGATCACTTGAATCCAGGAGTTTGAGA 392

Qy
332 CCAGGCTGGGCCAATACAGCGACTCTCTCTCTCTAAAA 370

Db
391 CCAGTCTGGGCAACACAGAGATACTCCGTCTCTACAAA 353

```

RESULT 8
US-10-027-632-150926
; Sequence 150926, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIORITY FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 150926
; LENGTH: 760
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-150926

```

```

1  APPLICANT: Wang, David G.
2  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
3  TITLE OF INVENTION: Polymorphisms in the Human Genome
4  FILE REFERENCE: 108827.129
5  CURRENT APPLICATION NUMBER: US/10/027,632
6  CURRENT FILING DATE: 2002-04-30
7  PRIOR APPLICATION NUMBER: US 60/218,006
8  PRIOR FILING DATE: 2000-07-12
9  PRIOR APPLICATION NUMBER: US 60/198,676
10 PRIOR FILING DATE: 2000-04-20

```

Db	272	GTAAAGTCATCTCAAGCTGGGCATGGTGGCTTGCACTCTTAATCCCAAGTGACITGGGAGG	213
Qy	294	CTGAGGCAGGAGGATTACTGGAGGCCAGGAGTTTGAGACCAGGCTGGGCAATACAGCGAG	353
Db	212	CTGAGGCAGGAGCATGCTTGGAGCCAGGAGTTTGAGACCAGCTTGGGCAACACAGCGAG	153
Qy	354	ACTCTCTCTCTCTAAAAA	370
Db	152	ACCTTGTCTCTGTAAAA	136

RESULT 11
 US-10-741-601-5691/c
 ; Sequence 5691, Application US/10741601
 ; Publication No. US20040166519A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: CL001500
 ; CURRENT FILING DATE: 2003-12-22
 ; NUMBER OF SEQ ID NOS: 26415
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 5691
 ; LENGTH: 67126
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(67126)
 ; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-10)
 US-10-741-601-5691

Query Match	23.8%	Score 88.2	DB 7	Length 67126
Best Local Similarity	76.6%	Pred. No. 1.3e-18		
Matches 108	Conservative 0	Mismatches 33	Indels 0	Gaps 0

Qy	230	ATATGAACCTCTCTCTGGCTAGGAGTGGCAGCCTACATCGCTAAATCCAGCAGCCTTGG	289
Db	16840	AAAAAAAAGTCACTCTGGCTGGAAGTGGTGGCTCAGGCTATATCCAGCAGCCTTGG	16781
Qy	290	GAAGCTGAGGCAGGAGGATTACTGGAGGCCAGGAGTTTGAGACCAGGCTGGGCAATACAG	349
Db	16780	GAAGCCAAGGCAGGTGGTCTTTGAGCCAGGAGTTTGAGACCAGCCTGGGCAACACAG	16721
Qy	350	CGAGACTCTCTCTCTAAAAA	370
Db	16720	TGAAACCCCTGTCTCTAAAAA	16700

RESULT 12
 US-10-741-600-17762/c
 ; Sequence 17762, Application US/10741600
 ; Publication No. US20050026169A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: CL001499
 ; CURRENT FILING DATE: 2003-12-22
 ; NUMBER OF SEQ ID NOS: 73997
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 17762
 ; LENGTH: 67126
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(67126)
 ; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-10)
 US-10-741-600-17762

```

US-10-741-600-17762

Query Match      23.8%; Score 88.2; DB 8; Length 67126;
Best Local Similarity 76.6%; Pred. No. 1.3e-18;
Matches 108; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy    230 ATATGAACTCTCCCTCGGTAGAGTGGCAGCCTACATCGCTAATCCCAGCACCTTGG 289
Db    | ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy    290 GAAGCTTGAGCGCAGGAGATTACTGGAGGCCAGGAGTTTCAGACCAGGCTGGGCATAACAG 349
Db    | ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy    16780 GARCCCAAGCAGGTGGTGGCTTGAGCCCGCAGGAGTTTGAGACCAGCCTGGGCAACAG 16721
Db    | ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy    350 CGAGACTCTCTCTCTAAAAA 370
Db    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy    16720 TGAACCCCTGCTCTCTAAAAA 16700
Db    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-09-925-065A-917415
; Sequence 917415, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108927.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 917415
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-917415

Query Match      23.6%; Score 87.4; DB 4; Length 707;
Best Local Similarity 74.1%; Pred. No. 2.5e-19;
Matches 109; Conservative 1; Mismatches 37; Indels 0; Gaps 0;

Qy    224 ACTCTCATGAACTCTCTCTCGGTAGGAGTGGCAGCCTACATCGCTAATCCAGCA 283
Db    | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
Qy    89  ACTCTGAAAAGAAAATTAACAAGCTGGGTGGCTCATCGCTATAATCCAGAA 148
Db    | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
Qy    284 CTTTGGGAGCTGAGCAGGAGGATTACTGGAGGCCAGGAGTTTCAGACCAGGCTGGGCA 343
Db    | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
Qy    149 CTTTGGGRGCTGAGCAGGAGGATCACTTGAGGCCAGGAGTTTGAGACCAGCCTGGGCA 208
Db    | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||

Qy    344 ATACAGCAGACTCTCTCTCTAAAAA 370
Db    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy    209 ACACAATGAGACTCTGCTCTAGAAAA 235
Db    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
US-10-027-632-325370
; Sequence 325370, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome

```

```

US-10-741-600-17762

Query Match      23.8%; Score 88.2; DB 8; Length 67126;
Best Local Similarity 76.6%; Pred. No. 1.3e-18;
Matches 108; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy    230 ATATGAACTCTCCCTCGGTAGAGTGGCAGCCTACATCGCTAATCCCAGCACCTTGG 289
Db    16840 AAAAAAAAAAAGTCATCTCGCTGGGAAGTGTTGGCTTCAGCCCTATAATCCCAGCACTTTGG 16781

Qy    290 GAAGCTTGAGCGCAGGAGGATTACTGGAGGCCAGGAGTTTTGAGACCAGGCTGGGCATATACAG 349
Db    16780 GARCCCAAGCAGGTGGTGGCTTGAGCCCGCAGGAGTTTGAGACCAGCCTGGGCAACACAG 16721

Qy    350 CGAGACTCTCTCTCTTA AAAA 370
Db    16720 TGAACCCCTGCTCTCTTA AAAA 16700


RESULT 13
US-09-925-065A-917415
; Sequence 917415, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108927.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 917415
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-917415

Query Match      23.6%; Score 87.4; DB 4; Length 707;
Best Local Similarity 74.1%; Pred. No. 2.5e-19;
Matches 109; Conservative 1; Mismatches 37; Indels 0; Gaps 0;

Qy    224 ACTCTCATATGAACCTCTCCTCTCGCTAGGAGTGGCAGCCTACATCGCTAATCCAGCA 283
Db    89   ACTCTGAAAAGAAGAAATTTAACAGGCTGGGTGGCTCATCGCTATAATCCAGAA 148

Qy    284 CTTTGGGAAGCTGAGCAGGAGGATTACTGGAGGCCAGGAGTTTGAGACCAGGCTGGGCA 343
Db    149 CTTTGGGRGCTGAGCAGGAGGATCACTTGAGGCCAGGAGTTTGAGACCAGCCTGGGCA 208

Qy    344 ATACAGCGAGACTCTCTCTCTTA AAAA 370
Db    209 ACACAATGAGACTCTGCTCTTAGAAAA 235


RESULT 14
US-10-027-632-325370
; Sequence 325370, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome

```

FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 325370
LENGTH: 847
TYPE: DNA
ORGANISM: Human
US-10-027-632-325370

Query Match 23.6%; Score 87.4; DB 5; Length 847;
Best Local Similarity 74.1%; Pred. No. 2.8e-19;
Matches 109; Conservative 1; Mismatches 37; Indels 0; Gaps 0;
QY 224 ACTCTCATATGAAACTCTCTCGCTAGGAGTGGCAGCGCTACATCGCTAAATCCAGCA 283
DB 140 ACTCTGAAAAGAAATATTAAACAGGCTGGTGGTGGCTCATGCTATATCCAGAA 199
QY 284 CTTTGGGAAGCTGAGGAGGAGGATTACTGGAGGCCAGGAGTTTGAGACCAGGCTGGGCA 343
DB 200 CTTTGGGRGGCTGAGGCGAGGAGATCATTGAGGCCAGGAGTTTGAGACCAGGCTGGGCA 259
QY 344 ATACAGCGAGACTCTCTCTTAAAAA 370
DB 260 ACACAATGAGACTCTGTCTTAGAAAA 286

RESULT 15
US-10-027-632-325370
Sequence 325370, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 325370
LENGTH: 847
TYPE: DNA
ORGANISM: Human

US-10-027-632-325370
Query Match 23.6%; Score 87.4; DB 6; Length 847;
Best Local Similarity 74.1%; Pred. No. 2.8e-19;
Matches 109; Conservative 1; Mismatches 37; Indels 0; Gaps 0;
QY 224 ACTCTCATATGAAACTCTCTCGCTAGGAGTGGCAGCGCTACATCGCTAAATCCAGCA 283
DB 140 ACTCTGAAAAGAAATATTAAACAGGCTGGTGGTGGCTCATGCTATATCCAGAA 199
QY 284 CTTTGGGAAGCTGAGGAGGAGGATTACTGGAGGCCAGGAGTTTGAGACCAGGCTGGGCA 343
DB 200 CTTTGGGRGGCTGAGGCGAGGAGATCATTGAGGCCAGGAGTTTGAGACCAGGCTGGGCA 259
QY 344 ATACAGCGAGACTCTCTCTTAAAAA 370
DB 260 ACACAATGAGACTCTGTCTTAGAAAA 286
Search completed: December 6, 2005, 22:56:44
Job time : 925.31 secs

QY	306	GATTACTGGAGGCCAGGAGTTTGAGACCCAGGCTGGGGCAATACAGCAGACTCTCTCTCT	365
Db	120315	GATTGCTTGAGCCAGGAGTTTAAGACCCAGGCTGGGGCAACATAGGAGATCTCTGTCTCTA	120374
QY	366	AAAAA 370	
Db	120375	CAAAA 120379	
RESULT 4			
US-10-496-711-25			
; Sequence 25, Application US/10496711			
; Publication No. US20050256649A1			
; GENERAL INFORMATION:			
; APPLICANT: SmithKline Beecham Corporation			
; TITLE OF INVENTION: HIGH THROUGHPUT CORRELATION OF			
; TITLE OF INVENTION: POLYMORPHIC FORMS WITH MULTIPLE PHENOTYPES WITHIN CLINICAL			
; TITLE OF INVENTION: POPULATIONS			
; FILE REFERENCE: PU4699WO			
; CURRENT APPLICATION NUMBER: US/10/496,711			
; CURRENT FILING DATE: 2004-05-26			
; PRIOR APPLICATION NUMBER: 60/344892			
; PRIOR FILING DATE: 2002-12-21			
; NUMBER OF SEQ ID NOS: 25			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 25			
; LENGTH: 7240			
; TYPE: DNA			
; ORGANISM: Homo Sapien			
US-10-496-711-25			
Query Match 21.4%; Score 79.2; DB 6; Length 7240;			
Best Local Similarity 77.4%; Pred. No. 5.4e-17;			
Matches 96; Conservative 0; Mismatches 28; Indels 0; Gaps 0;			
QY	247	TGGCTAGGAGTGGCAGCCTACATCGCTAATCCAGACACCTTTGGAGAGCTGAGCAGGAGG	306
Db	2357	TGGCTAGATGAGTGGCTCACACCTGTAATCCAGAACTTTGGAGAGCTGAGCAGG	2416
QY	307	ATTACTGGAGGCCAGGAGTTTGAGACCCAGGCTGGGGCAATACAGCAGACTCTCTCTCTA	366
Db	2417	ATTGCTTGAGGACAGGCATTAGACCCAGCCTGAGCAACATAGCAGACTCTCTCTCCAC	2476
QY	367	AAAA 370	
Db	2477	AAAA 2480	
RESULT 5			
US-10-276-233A-6			
; Sequence 6, Application US/10276233A			
; Publication No. US20050260572A1			
; GENERAL INFORMATION:			
; APPLICANT: DNA Chip Research Inc.			
; APPLICANT: Hitachi Software Engineering Co., Ltd.			
; TITLE OF INVENTION: A method of predicting cancer condition			
; FILE REFERENCE: PH-1533-PCT			
; CURRENT APPLICATION NUMBER: US/10/276,233A			
; CURRENT FILING DATE: 2002-11-14			
; PRIOR APPLICATION NUMBER: JP 2001-73063			
; PRIOR FILING DATE: 2001-03-14			
; PRIOR APPLICATION NUMBER: JP 2001-108503			
; PRIOR FILING DATE: 2001-04-06			
; PRIOR APPLICATION NUMBER: JP 2001-234807			
; PRIOR FILING DATE: 2001-08-02			
; NUMBER OF SEQ ID NOS: 27			
; SOFTWARE: PatentIn Ver. 3.2			
; SEQ ID NO 6			
; LENGTH: 79528			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-276-233A-6			

Query Match 21.4%; Score 79; DB 6; Length 79528;
Best Local Similarity 76.4%; Pred. No. 2e-16;
Matches 97; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 244 TCCTGGCTAGGAGTGGCAGCCTACATCGCTTAATCCAGCAGCTTGGGAAGCTGAGGAGG 303
Db 78708 TCTGGCCAGGCACTGTGGCTGCACACCTGTAAATCCAGCAGCTTGGGAGGCTTAGGCAGA 78767
|||
Qy 304 AGGATTACTGAGGCCAGGAGTTTGAGACACAGGCTGGGCAATACAGCGAGACTCTCTCTC 363
Db 78768 AGGATCGCTTGAGTCCAGGAGTTTGAGACACAGCCTGGGCAATATAGTGAGACTTTGTCTC 78827
|||
Qy 364 CTAATAA 370
Db 78828 AAAAAA 78834
|||

RESULT 6
US-11-112-908-45
; Sequence 45, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 45
; LENGTH: 182314
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-45

Query Match 21.3%; Score 78.8; DB 7; Length 182314;
Best Local Similarity 64.0%; Pred. No. 3.6e-16;
Matches 119; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy 185 CCGATTACAGACTTTCTTGATCTTTCTTCAGCATATGATCTCTCATATGAACTCTCTCT 244
Db 6606 CTGAATTAGAATATCAGTATGMACTCATGACATATTTCTTTAAATTAATAAATAAAT 6665
|||
Qy 245 CCTGCTAGGAGTGCCAGCCTACATCGCTTAATCCAGCAGCCTTGGGAAGCTGAGGCAGGA 304
Db 6666 CTAGGTCAGGTGTAGTGGCTCATGCTGTAAATCCAGCAGCTTTCAGATGCCAAGCAGGA 6725
|||
Qy 305 GGATTACTGGAGGCCAGAGTTTGAGACAGGCTGGGCAATACAGCGAGACTCTCTCTCC 364
Db 6726 GGAATGCTTGAGCCAGGAGTTGAGACAGCCTGGGCAATATAGGGAGACCCCTATCTCT. 6785
|||
Qy 365 TAAAAA 370
Db 6786 AGAAAA 6791
|||

RESULT 7
US-11-112-908-56/c
; Sequence 56, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; TITLE OF INVENTION: Breast Cancer Biomarkers

FILE REFERENCE: 04-164-US
CURRENT APPLICATION NUMBER: US/11/112,908
CURRENT FILING DATE: 2005-04-22
PRIOR APPLICATION NUMBER: US 60/564,758
PRIOR FILING DATE: 2004-04-23
PRIOR APPLICATION NUMBER: US 60/575,978
PRIOR FILING DATE: 2004-06-01
PRIOR APPLICATION NUMBER: US 60/631,702
PRIOR FILING DATE: 2004-11-30
PRIOR APPLICATION NUMBER: US 60/633,826
PRIOR FILING DATE: 2004-12-07
NUMBER OF SEQ ID NOS: 511
SOFTWARE: PatentIn version 3.3
SEQ ID NO 56
LENGTH: 150468
TYPE: DNA
ORGANISM: Homo sapiens
US-11-112-908-56

Query Match 21.2%; Score 78.4; DB 7; Length 150468;
Best Local Similarity 73.5%; Pred. No. 4.5e-16;
Matches 100; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 235 AAACCTCTCTCTGGCTAGGAGTGGCAGCCTACATCGCTTAATCCAGCAGCTTGGGAGGC 294
Db 45165 AAATTTAACTCAGGGCCAGCGCTGGTGGATCACACCTGCAATCCAGGACTTTGGGAGGC 45106
|||
Qy 295 TGAGGCAGGAGGATTACTGGAGCCAGGAGCTTTGAGACCAAGGCTGGGCAATACAGCGAGA 354
Db 45105 CGAGGCGGAGGATTGCTGAGGCCAGGAGTTTGAGACCAAGGCTGGGCAATACAGCGAGA 45046
|||
Qy 355 CTCTCTCTCTCTAAAAA 370
Db 45045 CTCTGGTCTCTACAAA 45030
|||

RESULT 8
US-11-112-908-55/c
; Sequence 55, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 55
; LENGTH: 193789
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-55

Query Match 21.2%; Score 78.4; DB 7; Length 193789;
Best Local Similarity 73.5%; Pred. No. 5.1e-16;
Matches 100; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 235 AAACCTCTCTCTGGCTAGGAGTGGCAGCCTACATCGCTTAATCCAGCAGCTTGGGAGGC 294
Db 95350 AAATTTAACTCAGGGCCAGCGCTGGTGGATCACACCTGCAATCCAGGACTTTGGGAGGC 95291
|||
Qy 295 TGAGGCAGGAGGATTACTGGAGGCCAGGAGTTTGAGACCAAGGCTGGGCAATACAGCGAGA 354
|||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 2600.1 Seconds
(without alignments)
6427.437 Million cell updates/sec

Title: US-09-980-046B-2
Perfect score: 294
Sequence: 1 cggccttaaggttcctctga.....aagagagtaattcccaaaa 294

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_hg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	286.6	97.5	163338	8	AL162426 Human DNA
2	286.6	97.5	172588	14	AC007430 Homo sapi
3	49.4	16.8	182870	2	AC116960 Dictyoste
4	49	16.7	110000	14	Continuation (13 o
5	49	16.7	110000	14	Continuation (14 o
6	48.8	16.6	105682	2	Continuation (4 of
7	48.6	16.5	155573	14	CR381703 Danio rer
8	48.4	16.5	188883	14	CR9333758 Danio rer
9	47.8	16.3	80518	14	AL109815 Plasmodiu
10	47.2	16.1	15714	6	AX346075 Sequence
11	47.2	16.1	15714	6	AX458542 Sequence
12	47	16.1	15714	5	BX322651 Zebrafish
13	47	16.0	152913	14	CR9333524 Danio rer
14	46.8	15.9	138046	14	AC117987 Felis cat
15	46.8	15.9	165105	5	AL929386 Zebrafish
16	46.4	15.8	110000	2	Continuation (5 of
17	46.4	15.8	160167	9	AC107810 Mus muscu
18	46.4	15.8	242741	14	AC106279 Rattus no

c 19	46.2	15.7	256879	2	AC116982
c 20	46	15.6	52838	14	AC087538 Homo sapi
c 21	46	15.6	93307	8	AL359737 Human DNA
c 22	45.8	15.6	24152	15	AY962591 Candida m
c 23	45.8	15.6	160491	14	CR387984 Danio rer
c 24	45.8	15.6	161532	8	AC146202 Pan trogl
c 25	45.8	15.6	188132	9	AC129290 Mus muscu
c 26	45.6	15.5	1321	15	AF120715 Saccharom
c 27	45.6	15.5	250421	2	AE014849 Plasmodiu
c 28	45.4	15.4	123205	9	AL935155 Mouse DNA
c 29	45.4	15.4	144449	14	AC067926 Mus muscu
c 30	45.4	15.4	162632	8	AC093916 Homo sapi
c 31	45.2	15.4	2664	2	AF206306 Plasmodiu
c 32	45.2	15.4	112714	5	BX511189 Zebrafish
c 33	45.2	15.4	153212	14	CR936952 Danio rer
c 34	45.2	15.4	270106	14	CR392327 Danio rer
c 35	44.8	15.2	14924	6	AX281279 Sequence
c 36	44.8	15.2	14924	6	AX345126 Sequence
c 37	44.8	15.2	110000	14	PFMAL8P1_04
c 38	44.8	15.2	176186	8	AC006925 Homo sapi
c 39	44.8	15.2	250029	2	AE014816 Plasmodiu
c 40	44.8	15.2	333321	2	AC116986 Dictyoste
c 41	44.6	15.2	1121	2	LMU00101 Leishmania
c 42	44.6	15.2	1141	2	KPLMDIVRA
c 43	44.6	15.2	123687	4	CR450381 Pig DNA s
c 44	44.6	15.2	154065	5	CR854838 Zebrafish
c 45	44.4	15.1	3488	2	AF447563 Plasmodiu

ALIGNMENTS

AL162426 163338 bp DNA linear PRI 18-MAY-2005
Human DNA sequence from clone Rp11-56D16 on chromosome 9 Contains
the STXBP1 gene for syntaxin binding protein 1
(UNC18, HUNC18, MUNC18-1), two novel genes, a novel gene (LOC138428),
the gene for a novel protein (LOC286207), the gene for a novel
protein (FLJ32780), the TOR2A gene for torin family 2, member A
(FLJ14771), the 3' end of the SH2D3C gene for SH2 domain containing
3C (CHAT.NSP3) and five CpG islands, complete sequence.

AL162426 GI:13624960
HTG; CHAT; CpG Island; FLJ14771; FLJ32780; HUNC18; LOC138428;
LOC286207; MUNC18-1; NSP3; SH2D3C; STXBP1; TOR2A; UNC18.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 163338)

Kimberley, A.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegasanger.ac.uk

Clone requests: clonerequest@sanger.ac.uk
On Apr 14, 2001 this sequence version replaced gi:13396434.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Information
on the WormPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
Chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr9
RP11-56D16 is from the library RPI-11.1 constructed by the group
of Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pBACe3.6

----- Genome Center
Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: vegas@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

source

1. 163338
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-56D16"
/clone_lib="RPC1-11.1"

misc_feature

1 /note="Clone_left_end: RP11-56D16"

gene

join(25578..25753,64916..64965,67028..67109,71688..71764,
73343..73421,74415..74518,76518..76666,78560..78644,
79479..79609,81393..81500,83211..83271,85364..85429,
86494..86574,89117..89255,89957..90066,91744..91845,
93470..93555,95719..95873,97681..97806,104088..106029)
/gene="STXBPl"

mRNA

/locus_tag="RP11-56D16.3-001"
join(25578..25753,64916..64965,67028..67109,71688..71764,
73343..73421,74415..74518,76518..76666,78560..78644,
79479..79609,81393..81500,83211..83271,85364..85429,
86494..86574,89117..89255,89957..90066,91744..91845,
93470..93555,95719..95873,97681..97806,104088..106029)
/gene="STXBPl"

/locus_tag="RP11-56D16.3-001"

/product="syntaxin binding protein 1"

/note="match: ESTs: AA040359.1 AA046287.1 AA053982.1
AA188032.1 AA325966.1 AA702633.1 AA984625.1 AI272340.1
AI372682.1 AI423921.1 AU151658.1 AU151659.1 AU523780.1
AU534938.1 AU562262.1 AU100377.1 AU125726.1 AU149329.1
AW869953.1 AW999802.1 BF097729.1 BE537233.1 BE789550.1
BQ206953.1 BF317196.1 BF348350.1 BF530933.1 BF923984.1
BF961379.1 BG023390.1 BG034484.1 BG473078.1 BG714378.1
BG760759.1 BG763316.1 BG764373.1 EG784677.1 BG766349.1
BG766441.1 BG768105.1 BI093878.1 BI258762.1 BI458224.1
BI459531.1 BI546893.1 BI547057.1 BI547741.1 BI552773.1
BI553312.1 BI553626.1 BI668003.1 BI753006.1 BI823666.1
BI861306.1 BI917323.1 BI963071.1 BM451362.1 BM453886.1
BM466332.1 BM547117.1 BM564692.1 BM652004.1 BM652915.1
BM673485.1 BM673893.1 BM685862.1 BM716079.1 BM717029.1
BM757624.1 BM804272.1 BM922153.1 BQ008989.1 BQ017105.1
BQ045480.1 BQ228860.1 BQ367373.1 BQ425660.1 BQ426010.1
BQ716065.1 BQ721161.1 BQ722040.1 BQ882009.1 BU075673.1
BU075862.1 BU620598.1 BU622583.1 BU633059.1 BU948274.1
BU948544.1 BU959179.1 C3939308.1 CA418034.1 H50918.1
H92032.1 RA3501.1 R85564.1 T30715.1 T33616.1 T66270.1
match: cDNAs: AF004562.1 BC031728.1 U21116.1"
join(25602..25753,64916..64965,67028..67109,71688..71764,
73343..73421,74415..74518,76518..76666,78560..78644,
79479..79609,81393..81500,83211..83271,85364..85429,
86494..86574,89117..89255,89957..90066,91744..91845,
93470..93555,95719..95873,104088..106029)
/gene="STXBPl"

/locus_tag="RP11-56D16.3-002"

/product="syntaxin binding protein 1"

join(25602..25753,64916..64965,67028..67109,71688..71764,
73343..73421,74415..74518,76518..76666,78560..78644,
79479..79609,81393..81500,83211..83271,85364..85429,
86494..86574,89117..89255,89957..90066,91744..91845,
93470..93555,95719..95873,104088..106029)
/gene="STXBPl"

/locus_tag="RP11-56D16.3-002"

/product="syntaxin binding protein 1"

/note="match: ESTs: AA324408.1 AA338739.1 AA883239.1

CDS

join(25717..25753,64916..64965,67028..67109,71688..71764,
73343..73421,74415..74518,76518..76666,78560..78644,
79479..79609,81393..81500,83211..83271,85364..85429,
86494..86574,89117..89255,89957..90066,91744..91845,
93470..93555,95719..95873,104088..104170)
/gene="STXBPl"

/locus_tag="RP11-56D16.3-002"

/standard_name="OTTHUMP0000022197"

/note="match: proteins: AAL37409 O08599 Q64320 Q8VD51"

/codon_start=1

/product="syntaxin binding protein 1"

/protein_id="CAI41181.1"

/db_xref="GI:57208972"

/db_xref="InterPro:IPR001619"

/translations="MAPIGLKAVVGKIMHDVTKVKKKGKVLVVDQLSMRLSSC
CMTDLMTEGITIVDINKREPLPSLEAVLITPSEKSVHSLISDFKDPPTAKYRAA
HVFTDSCDALFNELVKRAAKVKTLEINTAFLEPYEQVSLDSADFQSFYSPH
MGKPNPILRLAEQIATLCATKEYPAVRGEYKDNALLAQIQLKDAYKADDP
MGSPDKARSQILLDRGPPSPVHLHETLFQAMYSKMTGKTTMRDLSQLMKM
KEVLDDDLIALSHKHIAEYSQVTRSLKDFSSKRMNTGKTTMRDLSQLMKM
PQKLSKYSTHLHLAEDCMKHQGTVDKLCRVEQDLAMGTDAEGKIDPMRAIYP
ILLDANVYDKIRIILLYIFLKNGITENLNKLIQHAQIPPEDESEIITNHLGVPI
VTDTLRRSKPERKISQYQLSRWTPFIKIDMEDTIEDKLTDPHTYISTRSA
SFSTTAVSARYGHHNKNKAPGEYRSGPRLIIFILGGVSLNEMRCAYEVTOANGKWEVL
IGSTHILTPKGLDRLKLNKLTDEEISS"

join(25717..25753,64916..64965,67028..67109,71688..71764,
73343..73421,74415..74518,76518..76666,78560..78644,
79479..79609,81393..81500,83211..83271,85364..85429,
86494..86574,89117..89255,89957..90066,91744..91845,
93470..93555,95719..95873,97681..97790)
/gene="STXBPl"

/locus_tag="RP11-56D16.3-002"

/standard_name="OTTHUMP0000022196"

/note="match: proteins: AAH31728 Q967G8"

/codon_start=1

/product="syntaxin binding protein 1"

/protein_id="CAI41180.1"

/db_xref="GI:57208971"

/translations="MAPIGLKAVVGKIMHDVTKVKKKGKVLVVDQLSMRLSSC
CMTDLMTEGITIVDINKREPLPSLEAVLITPSEKSVHSLISDFKDPPTAKYRAA
HVFTDSCDALFNELVKRAAKVKTLEINTAFLEPYEQVSLDSADFQSFYSPH
MGKPNPILRLAEQIATLCATKEYPAVRGEYKDNALLAQIQLKDAYKADDP
MGSPDKARSQILLDRGPPSPVHLHETLFQAMYSKMTGKTTMRDLSQLMKM
KEVLDDDLIALSHKHIAEYSQVTRSLKDFSSKRMNTGKTTMRDLSQLMKM
PQKLSKYSTHLHLAEDCMKHQGTVDKLCRVEQDLAMGTDAEGKIDPMRAIYP
ILLDANVYDKIRIILLYIFLKNGITENLNKLIQHAQIPPEDESEIITNHLGVPI
VTDTLRRSKPERKISQYQLSRWTPFIKIDMEDTIEDKLTDPHTYISTRSA
SFSTTAVSARYGHHNKNKAPGEYRSGPRLIIFILGGVSLNEMRCAYEVTOANGKWEVL
IGSTHILTPKGLDRLHHPDRESSRVSPEDQAPTE"

25934..25965

/gene="STXBPl"

/locus_tag="RP11-56D16.3-002"

/note="Single clone region. Sequence from reads from a
short insert library derived from a single pUC clone.
Restriction digest data confirm the assembly"

39099..39123

/gene="STXBPl"

/locus_tag="RP11-56D16.3-002"

/note="Sequence from overlapping clone BA373J8 (AL390116).
Assembly confirmed by restriction digest"
71520

misc_feature

25934..25965

/gene="STXBPl"

/locus_tag="RP11-56D16.3-002"

/note="Single clone region. Sequence from reads from a
short insert library derived from a single pUC clone.
Restriction digest data confirm the assembly"

39099..39123

/gene="STXBPl"

/locus_tag="RP11-56D16.3-002"

/note="Sequence from overlapping clone BA373J8 (AL390116).
Assembly confirmed by restriction digest"
71520

misc_feature

25934..25965

/gene="STXBPl"

/locus_tag="RP11-56D16.3-002"

/note="Sequence from overlapping clone BA373J8 (AL390116).
Assembly confirmed by restriction digest"
71520


```
2.71, 5.62 - GSCJ_ID dd_00996"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO53118.1"
/db_xref="GI:28850333"
/translation="MELEKVEKDLLEVGVDPDIPSELISFGNESSWLNTNSDKF
KTFLN"
complement(join(9304..9432,9619..10184,10260..12186,
12335..12463))
/note="GeneID exon scores (in order of location ranges):
2.33, 62.67, 203.85, 8.24 - GSCJ_ID dd_00995"
/codon_start=1
/product="similar to Homo sapiens (Human). DNA mismatch
repair protein Msh2"
/protein_id="AAO53119.1"
/db_xref="GI:28850334"
/translation="MSDNEORESSOVVLKEDKTVTFPQSILVSNEDDTLRLPDKK
SLKWSIDPAPKKIKKIDNGLSTTTASSQQOQELGLAVLTIQGYEVENIVKEL
LDEKKKIEIWSKPNKSQOQWELIKGSGPNTQMFEDVLNGNCEGVMWALKVTRKG
SIVGIFSGDATFKTIGVSQFMNDNLNLSFIMQMSVKECLCCDQKNDYDYKVEK
KLSADGIPTELKPSFSKNAEQDLTLRLASVKNNLPDIQEHAIQASGLIKHLDL
LSPNPFKFLKDYLDORVMKLDSSFPKGLHIIDLKDSVSAAGAGGAGGAGSSSN
KDQSLYNLNOCTPMGSRLLLOWVKQPLNAEELRLNFEAFYNDELRQSLRSN
DLKIGDRLSKLHGKATLEDCLNLYGIVTRLPVVLQSLNHHSSHOBLIKVFI
ESLESIISDFAKFCAMVETIDLDLNDKHEYVIRKSFDETLRGILKKQKQISNKIER
FRVDTADLNDLDEAKVKLHYSEKQMFLLRIERKDEVAIRDKKYIVHATAKDGVRFAT
REIDTLNAYKKAAYELDQKDLAKRTLQIAASEVPLIEDLSLIATLDVFTLSHV
SSIAPIPTRPPIIPLGSDENGATVIGGRHPCVEIIONVNFANDLARGOSQFQ
ITGTMGSGKFPFIROGLIVLMAOIGCFVPAQKATIAVDCILSRVAGDSOLRGVS
TFMAELSTYILKVATKNSLIIIDELRGSTYDGLGLANGIARIYICNIGGFCFLPA
THFELTILSDLPKWNKLNHVSANTNMTFTLLYKVEQPCDSFGIHAVILLANFSPQ
VIENAKQAKALESEFNTLKQNHKFLIEEFKFNFNFSNDVEKLSVLNLSLKYSID
IN"
complement(join(12839..13000,13400..13435))
/note="GeneID exon scores (in order of location ranges):
2.76, -2.24 - GSCJ_ID dd_03346"
/codon_start=1
/product="similar to Plasmodium falciparum. Hypothetical
protein"
/protein_id="AAO53120.1"
/db_xref="GI:28850335"
/translation="MLSNGLKIDQEKKFFILLIINNKNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNN"
14458..15615
/note="GeneID exon scores (in order of location ranges):
104.42 - GSCJ_ID dd_00993"
/codon_start=1
/product="similar to Mus musculus (Mouse). Similar to
S-adenosylmethionine decarboxylase 1"
/protein_id="AAO08497.2"
/db_xref="GI:28850336"
/translation="WENNNDYKNRSTEIFNEADAGFVYQSNDSYSDGPEGPKKLDI
RGPISKSGVKSIAGSPKVLRTIDKEKQTVLDSARCTIISQTSNDHMSYVLSES
SLFVPRRAMIKTCGTITLHLIAKQVKGECGLEVEWFSRKNLQPKSQVFPHC
SFSDEYFLNKLFDQAYVMGDHNLNLYADFRKNPTLQRTQTEFEVMMHDLDET
VMKQFPRKGVSAWDTVNSGADLLPGSMIDDFDPCGYSMGLLNEFYWTIITP
ESHCSVSDPTVALADVNLQKLVNFKPGRFTAALVAEDGAPCGDPYAFDQVNP
SVAIQNTVHGGDGVVVSNNYQDLKKVFNDDLIQSIESTNIDLIQV"
complement(16030..17082)
/note="GeneID exon scores (in order of location ranges):
70.61 - GSCJ_ID dd_00992"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO08498.2"
/db_xref="GI:28850337"
/translation="MKCFILNIIKRNKINPNIIFKASYPHSPFNNSNIIRNFSSNNKI
NENNNKEDNMIIPSRSLSESLGAYLPNEEYEEVDFDEEYEDENDDDNNN
NNNNNIEKEEKEEKEVENQNLKEGIFKGGKKFQDMYIRIKENHLSNKHSDIFND
VGEQMSFEIERVEERLAKYQEKWEILDSMDQSDFIYKDNEKVDLYNKISENFKA
DTHFETPFIITQLENQLIKPENGSLSMSISDSAFNQTQSDDISRSIKIKORIQ
KOKLTNHRINGENLVDFSLYREDPFIILNNAQFLAKFNLSKSEVENLLKYTIPIIVK
FNTGKSGHWNVIYET"
join(18407..18484,18580..21294)
```

```
/note="GeneID exon scores (in order of location ranges):
7.28, 279.86 - GSCJ_ID dd_03068"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAO53121.1"
/db_xref="GI:28850338"
/translation="MDXYDTILWLRLLEENSKCADCTDSPRYMNTTYGTVCVSCG
AIHRELGNRVKSISSDXTQODIERLEKVGKNMADEIWLKSWSQOYPLPFPBDEKRV
RDFTKMYIEKKWIKEGIKOSDFISPTQSPGTPPTTSPKQOQOQPSAVSQKQT
TSPLRNGEKEKILTSKDSLESFESLSLGNLLNTNQQOQOQOQOQOQOQOQOQOQ
TSQOISPIFOABINQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ
SEKPNHOSYLNEKSTWNADSLGNSSFSYKIOVHPRTKSPSTSIYDQYKQOQOQ
QOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ
NQMGHANQMNQMNQMNQMNQMNQMNQMNQMNQMNQMNQMNQMNQMNQMNQMNQMN
NQLALVPVEQNNNNNNNNNNQVYPOQPOQOQOQOQOQOQOQOQOQOQOQOQOQOQ
IMNNNSNIENGPNFGISQDVHLKHQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ
PATTTTTIHTPENGPNLPSSESSSSGNPQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ
SSSPESGNLSDRDEEOKAWRRIOEKNIQIONDEIARKQLBEEBIEFKTRPRNTE
PSPRDRTSLQLOQSQPOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ
SSGNLILKRLSGRSKSYDYDDPFEINPRPLKSGGYDDYDSFMRKNKRIPGGGSNY
QPKISQOQDYQDNNNNNNNNNNYQOQPSQOQIVVHQSNYNNNNNNNNNNNANRYSNECD
YCGKDVGLSEMSYHKNTCEILRQBRCSMCKLVKHMQMGDHLQOQOQOQOQOQOQOQOQ
YCGKDVGLSEMSYHKNTCEILRQBRCSMCKLVKHMQMGDHLQOQOQOQOQOQOQOQOQ
```

Query Match	16.8%	Score 49.4	DB 2	Length 182870
Best Local Similarity	49.8%	Pred. No. 3.7	Mismatches 0	Gaps 0
Matches 125	Conservative 0			
QY	44	TTTGACGATCTCTACTTCAAATATTCTTCCAGTGCTTTTGTATCTTCTCAATGTTAT	103	
DB	58468	TTTGATCTTACACATAATAATCATGATTTTAAAGAAGTTCATATATTTCTAAATTTA	58527	
QY	104	AGTGAATATATATATATACATGCAGAAAAAGCATAAATCACTATTGAATACGAAATG	163	
DB	58528	GACTAAAAATCATTTTACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	58587	
QY	164	TGAACCTCTATGTAACCTCAACTCAGGCTGCCTACCGTTTTTAAATTTCTCTCAATCTTT	223	
DB	58588	ATAATAGAGATCTCAATTTAGTAATCTCAGCCCATTTTTTTTTTTTTTTTTTAAATCCCA	58647	
QY	224	AAAAACAATATCTATCTTTACTAAAGATGTTTTCAGAGTAAGATAAAATACAAAGAGTA	283	
DB	58648	AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA	58707	
QY	284	ATTTCCAAAAA 294		
DB	58708	AAAAATAAAAAA 58718		

Sequence split into 28 fragments LOCUS PFMAL13 Accession AL844509

Fragment Name	Begin	End
PFMAL13_00	1	110000
PFMAL13_01	100001	210000
PFMAL13_02	200001	310000
PFMAL13_03	300001	410000
PFMAL13_04	400001	510000
PFMAL13_05	500001	610000
PFMAL13_06	600001	710000
PFMAL13_07	700001	810000
PFMAL13_08	800001	910000
PFMAL13_09	900001	1010000
PFMAL13_10	1000001	1110000
PFMAL13_11	1100001	1210000
PFMAL13_12	1200001	1310000
PFMAL13_13	1300001	1410000
PFMAL13_14	1400001	1510000
PFMAL13_15	1500001	1610000
PFMAL13_16	1600001	1710000
PFMAL13_17	1700001	1810000
PFMAL13_18	1800001	1910000
PFMAL13_19	1900001	2010000

RESULT 4
PFMAL13_12
WPCOMMENT


```
COMMENT
On Mar 19, 2005 this sequence version replaced gi:46879102.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK72P23
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 154326 bases at least Q40
Consensus quality: 154612 bases at least Q30
Consensus quality: 154899 bases at least Q20
Insert size: 155273; sum-of-contigs
Insert size: 156492; 2.2% error; agarose-fp
Quality coverage: 10.08x in Q20 bases; sum-of-contigs Quality
coverage: 10.00x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 10164: contig of 10164 bp in length
* 10165 10264: gap of 100 bp
* 10265 33043: contig of 22779 bp in length
* 33044 33143: gap of 100 bp
* 33144 51452: contig of 18309 bp in length
* 51453 51552: gap of 100 bp
* 51553 155573: contig of 104021 bp in length.
-----
FEATURES
source
1..155573
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKSY-72P23"
/clone_lib="Daniokey"
1..10164
/note="assembly_fragment:00010
fragment_chain:1"
10265..33043
/note="assembly_fragment:00564
fragment_chain:1"
33144..51452
/note="assembly_fragment:00195.0"
51553..155573
/note="assembly_fragment:01026"

misc_feature
1..10164
/note="assembly_fragment:00010
fragment_chain:1"

misc_feature
10265..33043
/note="assembly_fragment:00564
fragment_chain:1"

misc_feature
33144..51452
/note="assembly_fragment:00195.0"

misc_feature
51553..155573
/note="assembly_fragment:01026"

ORIGIN
Query Match 16.5%; Score 48.6; DB 14; Length 155573;
Best Local Similarity 51.2%; Pred. No. 5.4;
Matches 105; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 51 ATCTCTTACTTCAAATATTCCTCAGTGGTTTGTATCTCTCAATGTTATAGTAAA 110
Db 155368 ATATATATATTTAAAAATTTTTTTTTTTTTTTTTTTTTTTTANNNTTTTNNAAAT 155427

QY 111 TATAATAATACATGCAGAAAAAGCATAAATCACTATTGATAACGAAATGTGAACCT 170
Db 155428 AAAAAAATAAAAAAATAAAAAAATAAATAAATAAATAAAAAAATAAAAAA 155487

QY 171 CTATGTAACTGCAACTCAGCGTCGCTTACCGTTTTTAAATTTCTCTTAATCTTTAAAAACA 230
Db 155488 AATAAAAAATAAAAAAATTTATTTTATATTTTTTTTTTATAATAAAAAAATAAAAAA 155547

QY 231 AATACATCTTTTACTAAAGATGTTT 255
Db 155548 AAAAAATATTTTTTTTAAATATTTTT 155572
```

```
RESULT 8
CR933758
LOCUS
DEFINITION
CR933758 188883 bp DNA linear HTG 01-JUN-2005
Danio rerio clone CH211-203E21, WORKING DRAFT SEQUENCE, 3 unordered
pieces.
ACCESSION
CR933758 GI:66863493
VERSION
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS
Danio rerio (zebrafish)
SOURCE
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 188883)
McLaren,S.
Direct Submission
Submitted (30-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jun 1, 2005 this sequence version replaced gi:58197316.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zC203E21
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 187341 bases at least Q40
Consensus quality: 187488 bases at least Q30
Consensus quality: 187715 bases at least Q20
Insert size: 188683; sum-of-contigs
Quality coverage: 9.20x in Q20 bases; sum-of-contigs Quality
coverage: 9.05x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 90306: contig of 90306 bp in length
* 90307 90406: gap of 100 bp
* 90407 98661: contig of 8255 bp in length
* 98662 98761: gap of 100 bp
* 98762 188883: contig of 90122 bp in length.
-----
FEATURES
source
1..188883
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-203E21"
/clone_lib="CHORI-211"
1..90306
/note="assembly_fragment:01266
fragment_chain:1
clone_end:SP6
vector_side:left"
90407..98661
/note="assembly_fragment:00022
fragment_chain:1"
98762..188883
/note="assembly_fragment:00108
fragment_chain:1"

misc_feature
1..90306
/note="assembly_fragment:01266
fragment_chain:1
clone_end:SP6
vector_side:left"
90407..98661
/note="assembly_fragment:00022
fragment_chain:1"
98762..188883
/note="assembly_fragment:00108
fragment_chain:1"

misc_feature
1..90306
/note="assembly_fragment:01266
fragment_chain:1
clone_end:SP6
vector_side:left"
90407..98661
/note="assembly_fragment:00022
fragment_chain:1"
98762..188883
/note="assembly_fragment:00108
fragment_chain:1"

ORIGIN
Query Match 16.5%; Score 48.4; DB 14; Length 188883;
```



```
Db 6485 CTTTCCAACTTAAATATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6426
Qy 151 AATAAGGAATGTGAACCTCTATGTAAGTCACTCAGGCTGCTACCGCTTTTAAATTT 210
Db 6425 AAAATCCCAAAAAAATATATAAAAAAAAAAAAAAAAAAACTCAAACTAACTA 6373
Qy 211 CTCTCTAATCTTTAAAAACAATACTATCTTTTACTAAAGATGTTTCAGAGTAAGATAAAA 270
Db 6372 TTCCCTCATATTAATACAAATACCATATTAACCTAAATTTAATTCACAATAACATCAA 6313
Qy 271 TACA 274
Db 6312 AAAA 6309

RESULT 11
AX458542/c
LOCUS AX458542 15714 bp DNA linear PAT 08-JUL-2002
DEFINITION Sequence 88 from Patent W00246454.
ACCESSION AX458542
VERSION AX458542.1 GI:21725206
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
1 Schacht, O.
AUTHORS Diagnosis of diseases associated with angiogenesis
TITLE Patent: WO 0246454-A 88 13-JUN-2002;
JOURNAL Epigenomics AG (DE)
FEATURES
source
1 .15714
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

ORIGIN
Query Match 16.1%; Score 47.2; DB 6; Length 15714;
Best Local Similarity 52.9%; Pred. No. 22;
Matches 129; Conservative 0; Mismatches 108; Indels 7; Gaps 1;

Qy 31 AAGAAGATCAGGATTTGAGCATCTCTTACTTCAAATTTATCTTCAGTGGTTTGTATC 90
Db 6545 ACGAAATCCAACTTCGCTACCTTTATTTCTCTTTATCTTCTCATCCCAATC 6486
Qy 91 TTCTCAATGTTTATAGTGAATATAATAATACATGCAGAAAAAAGCATAAATCAACTATG 150
Db 6485 CTTTCCAACTTAAATTTATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 6426
Qy 151 AATAAGGAATGTGAACCTCTATGTAAGTCACTCAGGCTGCTACCGCTTTTAAATTT 210
Db 6425 AAAATCCCAAAAAAATATATAAAAAAAAAAAAAAAAAAACTCAAACTAACTA 6373
Qy 211 CTCTCTAATCTTTAAAAACAATACTATCTTTTACTAAAGATGTTTCAGAGTAAGATAAAA 270
Db 6372 TTCCCTCATATTAATACAAATACCATATTAACCTAAATTTAATTCACAATAACATCAA 6313
Qy 271 TACA 274
Db 6312 AAAA 6309

RESULT 12
BX322651
LOCUS BX322651 152913 bp DNA linear VRT 29-JAN-2004
DEFINITION Zebrafish DNA sequence from clone DKEYP-32G10, complete sequence.
ACCESSION BX322651
VERSION BX322651.3 GI:41392279
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 152913)
Pelam, S.
Direct Submission
Submitted (14-JAN-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zf1sh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jan 29, 2004 this sequence version replaced gi:30387034.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zf1sh-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
Zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat
discovery system (Zhirong Bao and Sean Eddy, submitted), and those
beginning 'drr' were identified by Rick Waterman (Stephen Johnson
lab, WashU). For further information see
http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEYP-32G10
is from a Zebrafish BAC library
VECTOR: pIndigoBAC-5

FEATURES
source

Location/Qualifiers
1 .152913
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEYP-32G10"
/clone_lib="DanioKeyPilot"

ORIGIN

Query Match 16.0%; Score 47; DB 5; Length 152913;
Best Local Similarity 59.3%; Pred. No. 11;
Matches 80; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
Qy 43 ATTTGAGCATCTTCTTACTTCAAATTTATTTCTCCAGTGGTTTTTGTATCTTCTCAATGTA 102
Db 56862 ATTGACCAAGTGTGAAGCATCTCTTTTTCCTGTATCATCTTTCTCTCTTTATGTCA 56921
Qy 103 TAGTGAATATATAATACATGCAGAAAAAGCAATAATCAACTATTTCATTAACCAAAAT 162
Db 56922 CAGTTAATGTGAATAAACTTAGCAAAACAAACCAATATTATATATCATTAGCAAAAT 56981
Qy 163 GTGAACCTCTATGTA 177
Db 56982 GTAAACCTTTATGTA 56996

RESULT 13

COMMENT

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Sep 23, 2003 this sequence version replaced gi:33386428.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfsh-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'dr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_reio/fishmask.shtml CH211-69J13 is from a CHORI-211 BAC library
VECTOR: pTARBAC2.1.

FEATURES

source

1. 165105
Location/Qualifiers
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clones="CH211-69J13"
/clone_lib="CHORI-211"

ORIGIN

Query Match	15.9%	Score 46.8;	DB 5;	Length 165105;
Best Local Similarity	50.9%	Fred. No. 11;		
Matches 111;	Conservative 0;	Mismatches 107;	Indels 0;	Gaps 0;
Qy	62	CAAAATTATCTTCAGTGGTTTGTGTTCTCTCAATGTTAGTGAATAATAATAC	121	
Db	155302	CAAAATTAATAAGACTACAGTGTATATATACATGTTAAATGATAAACAAATATA	155243	
Qy	122	ATGCAGAAAAAGCATAAATCAACTATTGGAATAACGAAAAATGTGAACCTCTATGTA	181	
Db	155242	ATTTTGTACTAGGAAATATGAACCTTTTGGAAAAATAAGGACATAAATTC	155183	
Qy	182	CAACTCAGCGTCCTACCGTTTAAATTTCTCTCTAATCTTTAAACAAATACTATCTT	241	
Db	155182	ATATTTCATATTAAATAGCTTTTCATACATATCTCTACAAACGTTGACAAAAAGC	155123	
Qy	242	TACTAAAGATGTTTCAGAGTAGATATAAAATACAAAGAG	279	
Db	155122	TTTTTGTGTTAGATTGAAGATTAGGCGCTACAG	155085	

Search completed: December 7, 2005, 04:21:51
Job time : 2605.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 462.191 Seconds
(without alignments)
4239.411 Million cell updates/sec

Title: US-09-980-046B-2
Perfect score: 294
Sequence: 1 cggccttaaggttcctctga.....aagagagtaattccccaaaaa 294

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq 21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	294	100.0	294	4 AAC89733	AAC89733 Human gas
2	288.2	98.0	100445	13 ABD33179	ABD33179 Human can
C 3	47.2	16.1	15714	6 ABL33173	ABL33173 Human imm
C 4	47.2	16.1	15714	6 ABQ67058	ABQ67058 Human ang
C 5	44.8	15.2	14924	6 ABL32224	ABL32224 Human imm
C 6	44.8	15.2	14924	6 ABL54321	ABL54321 Chemical
C 7	44	15.0	5311	6 ABL33019	ABL33019 Human imm
C 8	43.8	14.9	6197	6 ABN80257	ABN80257 Human che
C 9	43.8	14.9	34548	6 ABL70604	ABL70604 Chemical
C 10	43.6	14.8	510	12 ACH78930	ACH78930 Human gen
C 11	43.4	14.8	5204	6 ABL32901	ABL32901 Human imm
C 12	43.4	14.8	8170	6 ABK28258	ABK28258 DNA trans
C 13	43.4	14.8	8423	6 ABL33407	ABL33407 Human imm
C 14	43.2	14.7	14551	6 ABL34584	ABL34584 Human met
C 15	43.2	14.7	14551	7 ADS99845	ADS99845 Bisulphit
C 16	43	14.6	9347	6 ABL49336	ABL49336 Human pol
C 17	43	14.6	110000	2 AAU21209_04	Continuation (5 of
C 18	42.6	14.5	5980	6 ABL33189	ABL33189 Human imm
C 19	42.6	14.5	6274	4 AAS46322	AAS46322 Tumour su

C 20	42.6	14.5	6274	6 ABK33952	ABK33952 Human DNA
C 21	42.6	14.5	6274	6 ABK31239	ABK31239 Signal tr
C 22	42.6	14.5	6274	8 ADA20355	ADA20355 Prostate
C 23	42.6	14.5	6274	8 ADA84162	ADA84162 Human ren
C 24	42.6	14.5	6274	10 ADB54250	ADB54250 Pretreate
C 25	42.6	14.5	6274	10 ADB54122	ADB54122 Pretreate
C 26	42.2	14.4	5163	6 ABL33249	ABL33249 Human imm
C 27	42.2	14.4	7008	4 AAS46531	AAS46531 Tumour su
C 28	42.2	14.4	9289	10 ADB84033	ADB84033 5' regula
C 29	41.8	14.2	8873	6 ABK31211	ABK31211 Signal tr
C 30	41.8	14.2	8873	6 ABL70174	ABL70174 Chemical
C 31	41.8	14.2	8873	6 AAS61125	AAS61125 Human gen
C 32	41.8	14.2	11178	6 ABK31172	ABK31172 Signal tr
C 33	41.8	14.2	11178	6 ABL70507	ABL70507 Chemical
C 34	41.8	14.2	11178	6 AAS61059	AAS61059 Human gen
C 35	41.6	14.1	6104	6 ABL33125	ABL33125 Human imm
C 36	41.6	14.1	7829	6 ABL33105	ABL33105 Human imm
C 37	41.6	14.1	7829	6 ABK31287	ABK31287 Signal tr
C 38	41.6	14.1	7829	6 ABL70266	ABL70266 Chemical
C 39	41.6	14.1	7829	6 AAS61191	AAS61191 Human gen
C 40	41.6	14.1	8056	8 ABZ10246	ABZ10246 Haematopo
C 41	41.6	14.1	16579	10 ADB54245	ADB54245 Pretreate
C 42	41.6	14.1	16579	10 ADE37772	ADE37772 Human che
C 43	41.6	14.1	16579	13 ADS89547	ADS89547 Oligonuc
C 44	41.4	14.1	3500	8 ACF62826	ACF62826 Colon can
C 45	41.4	14.1	6222	6 ABL32692	ABL32692 Human imm

ALIGNMENTS

RESULT 1
AAC89733
ID AAC89733 standard; cDNA; 294 BP.

AC AAC89733;

DT 12-MAR-2001 (first entry)

DE Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 2.

KW Human; cytostatic; immunomodulator; immunostimulant; vulnery;

KW anti-inflammatory; neuroprotective; antibacterial; gene therapy;

KW gastrointestinal inflammation; immune system disorder; genetic disorder;

KW cancer; autoimmune disorder; infection; wound healing; ss.

OS Homo sapiens.

PN WO200073324-A2.

PD 07-DEC-2000.

PF 01-JUN-2000; 2000WO-US015191.

PR 01-JUN-1999; 99US-0137058P.

PA (DIGI-) DIGITAL GENE TECHNOLOGIES INC.

PI Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Hasel KW;

PS WPI; 2001-061508/07.

PT New polynucleotides and polypeptides, useful in gene therapy and in

PT diagnosing a pathological condition, e.g. for modulating gene expression

PT in gastrointestinal inflammation, or for treating cancers or genetic

PS disorders.

PS Claim 1; Page 84; 108pp; English.

CC The present sequence is one of a number of isolated human polynucleotides

CC which are useful in gene therapy, and for diagnosing a pathological

CC condition or a susceptibility to it. In particular, the polynucleotides

CC are useful for modulating gene expression in gastrointestinal

CC inflammation. The polynucleotides are useful for chromosome
CC identification, controlling gene expression through triple helix
CC formation or antisense DNA or RNA, or identifying individuals from minute
CC biological samples using DNA-based identification techniques. The
CC polynucleotides can also be used as an alternative to restriction
CC fragment length polymorphism (RFLP), by determining the actual base-by-
CC base DNA sequences of selected portions of an individual's genome. The
CC polynucleotides may also be used as molecular weight markers on Southern
CC gels, as diagnostic probes for the presence of a specific mRNA, as a
CC probe to substract-out known sequences in the process of discovering novel
CC polynucleotides, or as an antigen to elicit an immune response. The
CC polypeptides are useful in diagnostic procedures to detect a disorder.
CC The polynucleotides and polypeptides are useful for preventing, treating
CC or ameliorating immune system disorders, genetic disorders, cancers, some
CC autoimmune disorders, or infections. The polynucleotides and polypeptides
CC are also useful for differentiating, proliferating or attracting cells,
CC leading to the regeneration of tissues, especially in wounds or burns.
CC The polypeptides and polynucleotides may also be used as a food additive
CC or preservative

XX Sequence 294 BP; 106 A; 52 C; 41 G; 95 T; 0 U; 0 Other;
SQ

Query Match 100.0%; Score 294; DB 4; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.2e-56;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGCCTTAAGGTTCTCTGACAGTTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 60
Db 1 CGGCCTTAAGGTTCTCTGACAGTTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 60
Qy 61 TCAATATTATCTTCCAGTGGTTTGTATCTCTCAATGTATATGTAATATAATA 120
Db 61 TCAATATTATCTTCCAGTGGTTTGTATCTCTCAATGTATATGTAATATAATA 120
Qy 121 CATGAGAAAAAGCATAAATCAACTATTGAATAACGAAATGTGAACCTCTATGTAAT 180
Db 121 CATGAGAAAAAGCATAAATCAACTATTGAATAACGAAATGTGAACCTCTATGTAAT 180
Qy 181 GCAACTCAGGCTGCTACCGTTTTTAATTTCTCTAATCTTTAAAAACAATACTATCT 240
Db 181 GCAACTCAGGCTGCTACCGTTTTTAATTTCTCTAATCTTTAAAAACAATACTATCT 240
Qy 241 TTACTAAAGATGTTTCAGAGTAAGTAATAATACAAAGAGAGTAATTTCCCAAAA 294
Db 241 TTACTAAAGATGTTTCAGAGTAAGTAATAATACAAAGAGAGTAATTTCCCAAAA 294

RESULT 2
ABD33179
ID ABD33179 standard; DNA; 100445 BP.

XX ABD33179;

XX 18-NOV-2004 (first entry)

XX Human cancer-associated (CA) gene HD07-026.

XX Human; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
KW db; cancer; cytostatic.

XX Homo sapiens.

XX WO2004058146-A2.

XX 15-JUL-2004.

XX 15-DEC-2003; 2003WO-US040081.

XX 17-DEC-2002; 2002US-003228B1.

XX (SAGR-) SAGRES DISCOVERY INC.

XX Morris DW, Malandro MS;

XX

DR WPI; 2004-499109/47.

XX Novel human cancer associated protein encoded within open reading frame
PT of cancer associated gene, useful as targets for diagnosing cancer.

XX Claim 16; SEQ ID NO 170; 182pp; English.

XX The invention relates to cancer-associated proteins (CAP) and the cancer-
CC associated (CA) nucleic acids encoding them. The invention also relates
CC to a method for treating cancers involving administering to a patient an
CC inhibitor of CAP, and a method of screening for anticancer activity in a
CC potential drug involving providing a cell that expresses a CA gene,
CC contacting a tissue sample derived from a cancer cell with an anticancer
CC drug candidate and monitoring the effect of the anticancer drug candidate
CC on expression of the CA gene. The CAP proteins are useful for detecting
CC cancer associated with expression of a CAP protein in a test cell sample
CC and for screening for a bioactive agent capable of modulating the
CC activity of a CAP protein. The CA nucleic acids are useful for diagnosing
CC cancer, involving determining the expression of a CA nucleic acid in a
CC tissue. This sequence represents a human CA gene of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 100445 BP; 25647 A; 23783 C; 23245 G; 27750 T; 0 U; 20 Other;

Query Match 98.0%; Score 288.2; DB 13; Length 100445;
Best Local Similarity 99.0%; Pred. No. 5.1e-55;
Matches 290; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGGCCTTAAGGTTCTCTGACAGTTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 60
Db 61160 CGGCCTTAAGGTTCTCTGACAGTTGTCTCAAGAAGATCAGGATTTGAGCATCTCTTACT 61219
Qy 61 TCAATATTATCTTCCAGTGGTTTGTATCTCTCAATGTATAGTGAATAATAATA 120
Db 61220 TCAATATTATCTTCCAGTGGTTTGTATCTCTCAATGTATAGTGAATAATAATA 61279
Qy 121 CATGAGAAAAAGCATAAATCAACTATTGAATAACGAAATGTGAACCTCTATGTAAT 180
Db 61280 CATGAGAAAAAGCATAAATCAACTATTGAATAACGAAATGTGAACCTCTATGTAAT 61339
Qy 181 GCAACTCAGGCTGCTACCGTTTTTAATTTCTCTAATCTTTAAAAACAATACTATCT 240
Db 61340 GCAACTCAGGCTGCTACCGTTTTTAATTTCTCTAATCTTTAAAAACAATACTATCT 61399
Qy 241 TTACTAAAGATGTTTCAGAGTAAGTAATAATAACAAAGAGAGTAATTTCCCAAAA 293
Db 61400 TTACTAAAGATGTTTCAGAGTAAGTAATAATAACAAAGAGAGTAATTTCCCAAAA 61452

RESULT 3
ABL33173/c

ID ABL33173 standard; DNA; 15714 BP.

XX ABL33173;

XX 26-MAR-2002 (first entry)

XX Human immune system associated gene SEQ ID NO: 1146.

XX Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.

XX Homo sapiens.

XX OS

KW antiarteriosclerotic; antianemic; cytotatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antithyroid; antidiabetic; antipeptidic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.
XX Homo sapiens.
XX WO200200928-A2.
XX 03-JAN-2002.
XX 02-JUL-2001; 2001WO-EP007537.
XX 30-JUN-2000; 2000DE-01032529.
XX 01-SEP-2000; 2000DE-01043826.
XX (EPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX Claim 1; SEQ ID NO 197; 32pp + Sequence Listing; German.
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SQ Sequence 14924 BP; 4130 A; 179 C; 3007 G; 7602 T; 0 U; 6 Other;
Query Match 15.2%; Score 44.8; DB 6; Length 14924;
Best Local Similarity 48.1%; Pred. No. 1.8;
Matches 127; Conservative 0; Mismatches 137; Indels 0; Gaps 0;
Qy 31 AAGAAGATCAGGATTTGAGCATCTCTTACTTCAAAATATTCTCCAGTGGTTTGTATC 90
Db 12072 AATAAATAAATACTTTCTCTCTTCAAAAATATTATTTCAAATATATTTCTTTAA 12013
Qy 91 TTCTCAATGTTATAGTGAATATATAATATACATGCAGAAAAAGCATAAATCAACTATTG 150
Db 12012 TAATACTATTTATATCTTTAAACCATTTATATAAAATATAATATATAATCAAT 11953
Qy 151 AATAACGAAATGTGAACCTCTATGTAACCTCAGGCTGCCTACCGTTTAAATTT 210
Db 11952 ATCTATCAATATATAATATAATAATAATAATAATAATAATAATAATAATAATA 11893
Qy 211 CTCTCTAATCTTTAAAAACAATACTATCTTTACTAAAGATGTTTCAGAGTAAGATAAA 270
Db 11892 ATAATAATTTTAAATTTCAACTACCTTTAATACAAAAAACCTTAAACATAATA 11833
Qy 271 TACAAAGAGAGTAATTTCCCAAAA 294
Db 11832 TACAATCAATAATCCAAAAAA 11809
RESULT 6
ABL54321/c
ID ABL54321 standard; DNA; 14924 BP.
XX
AC ABL54321;
XX

DT 29-JUL-2002 (first entry)
XX Chemically treated apoptosis gene #11.
XX Apoptosis; HIV; Bloom syndrome; cardiopathy; neurodegenerative disorder;
KW Herpes simplex virus; renal ischaemia; amyotrophic lateral sclerosis;
KW cancer; ds.
XX Unidentified.
XX WO200177164-A2.
XX 18-OCT-2001.
XX 06-APR-2001; 2001WO-EP003969.
XX 06-APR-2000; 2000DE-01019058.
XX 07-APR-2000; 2000DE-01019173.
XX 30-JUN-2000; 2000DE-01032529.
XX 01-SEP-2000; 2000DE-01043826.
XX (EPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-017444/02.
XX Chemically modified sequences of genes associated with apoptosis are
PT useful to determine methylation patterns of genomic DNA samples for
PT diagnosis of associated diseases such as cancer.
XX
PS Claim 1; Seq ID #21; 24pp; English.
XX This invention relates to chemically pre-treated DNA of genes associated
CC with apoptosis. The nucleic acids are used to allocate patients for
CC specific therapy for HIV infection, Bloom syndrome, cardiopathy, aging,
CC neurodegenerative disorders, Herpes simplex virus infection, renal
CC ischaemia, amyotrophic lateral sclerosis, solid tumours and cancers. This
CC nucleotide sequence represents a chemically treated apoptosis gene. Even
CC SEQ ID numbers are the complementary DNA strands to the odd SEQ ID
CC numbers. The sequence data for this patent is not represented in the
CC printed specification but is based on information supplied by the
CC European patent office
XX
SQ Sequence 14924 BP; 4130 A; 179 C; 3007 G; 7602 T; 0 U; 6 Other;
Query Match 15.2%; Score 44.8; DB 6; Length 14924;
Best Local Similarity 48.1%; Pred. No. 1.8;
Matches 127; Conservative 0; Mismatches 137; Indels 0; Gaps 0;
Qy 31 AAGAAGATCAGGATTTGAGCATCTCTTACTTCAAAATATTCTCCAGTGGTTTGTATC 90
Db 12072 AATAAATAAATACTTTCTCTCTTCAAAAATATTATTTCAAATATATTTCTTTAA 12013
Qy 91 TTCTCAATGTTATAGTGAATATATAATATACATGCAGAAAAAGCATAAATCAACTATTG 150
Db 12012 TAATACTATTTATATCTTTAAACCATTTATATAAAATATAATATATAATCAAT 11953
Qy 151 AATAACGAAATGTGAACCTCTATGTAACCTCAGGCTGCCTACCGTTTAAATTT 210
Db 11952 ATCTATCAATATATAATATAATAATAATAATAATAATAATAATAATAATAATA 11893
Qy 211 CTCTCTAATCTTTAAAAACAATACTATCTTTACTAAAGATGTTTCAGAGTAAGATAAA 270
Db 11892 ATAATAATTTTAAATTTCAACTACCTTTAATACAAAAAACCTTAAACATAATA 11833
Qy 271 TACAAAGAGAGTAATTTCCCAAAA 294
Db 11832 TACAATCAATATCCAAAAAA 11809
RESULT 7
ABL33019/c

CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704

XX
SQ Sequence 510 BP; 158 A; 79 C; 69 G; 204 T; 0 U; 0 Other;

Query Match 14.8%; Score 43.6; DB 12; Length 510;
Best Local Similarity 48.8%; Pred. No. 2.2;
Matches 118; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 2 GGCTTAAAGTTCCTCTGACAGTGTCTCAAGAAGATCAGGATTTTGAGCATCTCTTACTT 61
DB 225 GGCACTAAGTCTCTCTGATGGTTCCTTTTCAGTCTTACTTTATTTAGATCTTCTT 284
QY 62 CAAATTAATTTCTCCAGTGGTTTGTGTTATCTTCTCAATGTTATAGTGAATATAATAATAC 121
DB 285 CAGACTCTGAATTAAGTCTCTGGTTGTAGTCAAGATGCAATTTGATCTTTTAAATAATGGA 344
QY 122 ATGCAGAAAAGCAATAATCAACTATTGAAACGAAATGTGAACCTCTATGTAATG 181
DB 345 GGAAGAAGAAGATTAATAATATATAGACATTTTTCAAAAGTTTGAAGTCCAAATGTAAT 404
QY 182 CAACTCAGGCTGCTACCGTTTTTAAATTTCTCTCTAATCTTTTAAACAAATACTATCT 241
DB 405 TAGTTATAGTCTTTTAAATTAATCTTTTTTTTTGAAATCTTAAATAATCCCATACTT 464
QY 242 TA 243
DB 465 CA 466

RESULT 11

ABL32901/c
ID ABL32901 standard; DNA; 5204 BP.

XX
AC ABL32901;

XX
DT 26-MAR-2002 (first entry)

XX
DE Human immune system associated gene SEQ ID NO: 874.

XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cyostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
DB.

XX
OS Homo sapiens.

XX
PN WO200200928-A2.

XX
PD 03-JAN-2002.

XX
PF 02-JUL-2001; 2001WO-EP007537.

XX
PR 30-JUN-2000; 2000DE-01032529.

XX
PR 01-SEP-2000; 2000DE-01043826.

XX
PA (EPIG-) EPIGENOMICS AG.

XX
PI Olek A, Piepenbrock C, Berlin K;

XX
DR WPI; 2002-130909/17.

XX
PT Nucleic acid comprising fragment of chemically modified gene, useful for

PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.

XX
PS Claim 1; SEQ ID NO 874; 32pp + Sequence Listing; German.

XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention

XX
SQ Sequence 5204 BP; 1463 A; 72 C; 1014 G; 2655 T; 0 U; 0 Other;

Query Match 14.8%; Score 43.4; DB 6; Length 5204;
Best Local Similarity 50.2%; Pred. No. 3.2;
Matches 107; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 82 TTTGTTATCTTCTCAATGTTATAGTGAATATAATAATACATGCAGAAAAAGCAATAAT 141
DB 2290 TTTTTCACAGTAAACGCCCTAATAACAAAAAATAAAAAAATAAAAAA 2231
QY 142 CAACTATTGAATAACGAAATGTGAACCTCTATGTAACTGCAACTCAGCTCCCTACCGT 201
DB 2230 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 2171
QY 202 TTTTAAATTTCTCTTAATCTTTTAAACAAATACTATCTTTTAAAGATGTTTCAGACT 261
DB 2170 TAAATTTCAATTTCTTCAAAATCCTATACATATAATTTTAAACCGTTCTACAACT 2111
QY 262 AAGATAAAATACAAAGAGAGATTAATTCCTCAAAAA 294
DB 2110 AAAAAAATAATACTAATAATAACCCCAATA 2078

RESULT 12

ABK28258/c

ID ABK28258 standard; DNA; 8170 BP.

XX
AC ABK28258;

XX
DT 23-APR-2002 (first entry)

XX
DE DNA transcription associated complementary genomic DNA #66.

XX
KW DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;
KW PNA; cytosine methylation state; SNP; retroviral infection; gene; ds;
KW single nucleotide polymorphism; adenosine deaminase deficiency; cancer;
KW viral infection; Sezary syndrome; haematological disorder; tuberculosis;
KW immunological disorder; Werner syndrome; developmental disorder;
KW psoriasis; Rieger's syndrome; neurological disorder; erythropeleia;
KW neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease;
KW myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;
KW angiogenesis; congenital heart disease; HDR syndrome; gene therapy;
KW polyglutamine disorder; solid tumour.

XX
OS Unidentified.

XX
PN WO200192565-A2.

XX
PD 06-DEC-2001.

XX
PF 06-APR-2001; 2001WO-EP003973.

XX
PR 06-APR-2000; 2000DE-01019058.

XX
PR 07-APR-2000; 2000DE-01019173.

XX
PR 30-JUN-2000; 2000DE-01032529.

XX
PR 01-SEP-2000; 2000DE-01043826.

XX
XX (EPIG-) EPIGENOMICS AG.

XX
PA

```
PI Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-090046/12.
XX
XX New nucleic acids or oligomers, useful for diagnosing or treating
XX diseases associated with DNA transcription, e.g. immunological disorders,
XX Werner syndrome, psoriasis, myocardial infarction, solid tumors or
XX cancer.
XX
XX Claim 1; SEQ ID NO 132; 32pp; English.
XX
XX The invention relates to a nucleic acid, which comprises a segment of the
XX chemically pretreated DNA of genes associated with DNA transcription from
XX one of 346 sequences, and an oligomer, in particular an oligonucleotide
XX or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical
XX to the chemically pretreated DNA of genes associated with DNA
XX transcription. The set of oligomer probes are useful for detecting the
XX cytosine methylation state and/or single nucleotide polymorphisms (SNPs)
XX in a chemically pretreated genomic DNA. The nucleic acids are useful for
XX diagnosing or treating diseases associated with DNA transcription
XX (particularly with the methylation status), e.g. adenosine deaminase
XX deficiency, viral infection, retroviral infection, Sezary syndrome,
XX haematological disorders, immunological disorders, Werner syndrome,
XX tuberculosis, developmental disorders, psoriasis, Rieger's syndrome,
XX neurological disorders, neurodegenerative disorders, Waardenburg
XX syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial
XX infarction, hypertension, angiogenesis, erythropoiesis, congenital heart
XX diseases, HDR syndrome, arthritis, polyglutamine disorders, solid tumours
XX or cancer. Sequences ABK28127-ABK28472 represent DNA transcription
XX associated genomic DNA molecules of the invention. Note: The sequence
XX data for this patent did not form part of the printed specification but
XX was obtained in electronic format directly from the European Patent
XX Office
XX
XX Sequence 8170 BP; 2790 A; 69 C; 1334 G; 3977 T; 0 U; 0 Other;

Query Match 14.8%; Score 43.4; DB 6; Length 8170;
Best Local Similarity 50.2%; Pred. No. 3.4;
Matches 107; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Qy 82 TTGTGTTCTCTCAATGTTATAGTGAATATATAATACATGAGAAAAAGCATAAAT 141
Db TTTTAAATCTAAACACCTTTTATATAAATAATTTATATAATATAAACAATATAATAA 2569

Qy 142 CAACATTATGAATAACGAAATGTGAACCTCTATGTAACCTGCAACTGCGCTACCGT 201
Db ATATTATAAATAACAATATTAAAAAATACTACTTCTTTATCTTAAAAATTTATTCGAT 2509

Qy 202 TTTTAATTTCTCTCTAAATCTTTTAAAAACAATACTATCTTTTACTAAAGATGTTTCAGAGT 261
Db TATTAAATTTACATTTTACTATCTTAAAAACAACCTATATAACCTTATATTAAACAACAAC 2449

Qy 262 AAGATAAAATACAAAGAGAGTAATTTCCCAAAA 294
Db AATCTAAATATACAAACAACATAAATTTTCATAAA 2416

RESULT 13
ABL33407/c
ID ABL33407 standard; DNA; 8423 BP.
XX
XX ABL33407;
XX
XX 26-MAR-2002 (first entry)
XX
XX Human immune system associated gene SEQ ID NO: 1380.
XX
XX Human; immune system disease; cytosine methylation; antiasthmatic;
XX antiarteriosclerotic; antianaemic; cytostatic; nootropic;
XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
XX antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
XX antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;

KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.
XX Homo sapiens.
XX
XX WO200200928-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP007537.
XX
XX 30-JUN-2000; 2000DE-01032529.
XX
XX 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful for
XX diagnosis and treatment of diseases associated with abnormal cytosine
XX methylation.
XX
XX Claim 1; SEQ ID NO 1380; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX diseases. The present sequence is a gene of the invention
XX
XX Sequence 8423 BP; 2529 A; 214 C; 1834 G; 3845 T; 0 U; 1 Other;

Query Match 14.8%; Score 43.4; DB 6; Length 8423;
Best Local Similarity 51.9%; Pred. No. 3.4;
Matches 98; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy 86 TTATCTTCTCAATGTTATAGTGAATATAATAATACATGCAAAAAAGCAATAAATCAAC 145
Db TCATCCCAAAACATAACCTCCCAATATATCTCTCAATATATAAACAATCAATAAATAAC 1115

Qy 146 TATTGAATAACGAAATGTGAACCTCTATGTAACCTGCAACTCAGGCTGCGCTTTT 205
Db AAAAAACATTTTAAAAATAAAAACTTTAATTTAAATTAATAATCAATAAAAAAACTCTT 1055

Qy 206 AATTTCTCTCTAAATCTTTAAAAACAATACTATCTTTTACTAAAGATGTTTCAGAGTAGA 265
Db ACTTCTCCAAATATTAAAAACAACAATAACGATATATAAAAAAATAAATAAATTTTATA 995

Qy 266 TAAATACA 274
Db TAAATTACA 986

RESULT 14
ABL34584/c
ID ABL34584 standard; DNA; 14551 BP.
XX
XX ABL34584;
XX
XX 26-MAR-2002 (first entry)
XX
XX Human metastasis associated gene SEQ ID NO: 137.
XX
XX Metastasis associated gene; cytostatic; gene therapy; cancer;
XX cytosine methylation; gene; ds.
XX
XX Homo sapiens.
XX
```

```
PN WO200177376-A2.
XX
XX
PD 18-OCT-2001.
XX
XX
PF 06-APR-2001; 2001WO-EP003970.
XX
XX 06-APR-2000; 2000DE-01019058.
XX 07-APR-2000; 2000DE-01019173.
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2002-010922/01.
DR
XX New nucleic acid derived from chemically treated metastasis genes, useful
PT for diagnosis of cancers by analysis of cytosine methylation, also for
PT treatment.
XX
XX Claim 1; SEQ ID NO 137; 23pp + Sequence Listing; English.
PS
XX The present invention provides a number of human metastasis associated
CC genes which are modified by cytosine methylation. The sequences can be
CC used in the diagnosis and treatment of cancer. The present sequence is
CC one of the genes of the invention
XX
XX Sequence 14551 BP; 4362 A; 107 C; 3194 G; 6888 T; 0 U; 0 Other;
SQ
Query Match 14.7%; Score 43.2; DB 6; Length 14551;
Best Local Similarity 52.7%; Pred. No. 4.1;
Matches 117; Conservative 0; Mismatches 103; Indels 2; Gaps 1;
OY 57 TACTTCAAATATTCTTCCAGTGGTTTGTATCTTCTCAATGTTATAGTGAATATAAT 116
Db 1078 TAATACTAAACACTATATAAAAAATTTTCTTATTTAATACATCTCAAACTAAATAAAAA 1019
OY 117 AATACATGCGAAGAAACATATACTATTGTAATA--CGAAATGTGAACCTCTAT 174
Db 1018 CATACATTAATAATAAAATCAAAATAATTTAATAAAAAATTTCTAAATTTCTAAATAAAT 959
OY 175 GAACTGCAATCAGGCTGCTTACCGTTTCTTAAATTTTCTCTTAATCTTTAAAAACAATA 234
Db 958 TCTATATAAAATCTCTTACCCCAAAATCTTCTTCACTAAACGTCGAAATATAAC 899
OY 235 CTATCTTTACTAAAGATGTTTCAGAGTAAGATAAAATACAA 276
Db 898 TTATTTAAATCAAAATCTTTCAAATATAATAATAAAAA 857
RESULT 15
ADS99845/c
ID ADS99845 standard; DNA; 14551 BP.
XX
XX ADS99845;
XX
XX 02-DEC-2004 (first entry)
XX
XX Bisulphite treated human gene associated with metastasis #69.
XX
XX Human; ds; gene; Bisulphite; metastasis; cancer; cytosine;
KW DNA methylation; matrix-assisted laser desorption/ionisation; MALDI;
KW electrospray; mass spectrometry; CpG dinucleotide; solid tumour.
XX
XX Homo sapiens.
XX
XX US2003148327-A1.
XX
XX 07-AUG-2003.
XX
XX 21-JAN-2003; 2003US-00240485.
XX
```

```
PR 06-APR-2000; 2000DE-01019058.
PR 07-APR-2000; 2000DE-01019173.
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
PR 06-APR-2001; 2001WO-EP003970.
XX
XX (OLEK/) OLEK A.
PA (PIEP/) PIEPENBROCK C.
PA (BERL/) BERLIN K.
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2002-010922/01.
DR
XX New nucleic acid derived from chemically treated metastasis genes, useful
PT for diagnosis of cancers by analysis of cytosine methylation, also for
PT treatment.
XX
XX Claim 1; SEQ ID NO 137; 9pp; English.
PS
XX The invention relates to a nucleic acid comprising at least 18 bases from
CC a segment of the chemically pretreated DNA of genes associated with
CC metastasis, i.e. any of ADS9709-ADS9906 human genomic sequences or any
CC of the 19 sequences appearing as ADS9911-ADS9929. SEQ ID 2,4,6 etc are
CC the complements of SEQ ID 1,3,5, etc. Also included are an oligomer
CC (particularly an oligonucleotide or peptide nucleic acid) comprising at
CC least one base sequence of at least 9 bases which hybridises to (or is
CC identical with) the sequences referred to above, producing an array of
CC the oligomers on a carrier, obtaining genetic and/or epigenetic
CC parameters for diagnosis and/or therapy of diseases (or predisposition to
CC them) by analysis of cytosine methylation and a kit comprising a
CC bisulphite (disulphite or hydrogen sulphite) and the oligomers. In the
CC method of above 5-unmethylated cytosines in a genomic DNA sample are
CC converted chemically to uracil, or another base with hybridisation
CC properties different from those of cytosine, then fragments of the
CC treated DNA amplified (particularly by polymerase chain reaction) using
CC the oligomers and a polymerase (preferably heat stable) to produce
CC labelled amplicons. These are tested for hybridisation to an array of
CC oligomers and any hybridisation detected. The amplicons are labelled with
CC fluorescent or radioactive markers, or with a detachable mass marker to
CC allow their detection by mass spectrometry, specifically using the matrix
CC assisted laser desorption/ionisation (MALDI) or electrospray techniques.
CC To improve detection in the mass spectrometer, fragments formed in the
CC instrument have only a single net charge (positive or negative). The
CC genomic DNA is from e.g. a cell line, biopsy sample, blood, or paraffin-
CC embedded tissue sample. Oligonucleotides or peptide-nucleic acids that
CC are complementary to (or identical with) parts of the nucleic acids listed
CC above may be used as primers for amplification of the nucleic acids or
CC their complements, and for determining cytosine methylation status and/or
CC single nucleotide polymorphisms in metastasis-related genes. They can be
CC used for analysis of diseases associated with methylation of CpG
CC dinucleotides and to determine (epigenetic parameters for diagnosis
CC and/or therapy of disease (or predisposition)). The genomic DNA sequences
CC are useful for diagnosis and therapy of solid tumours and cancer. The
CC present sequence is a bisulphite treated human gene associated with
CC metastasis. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html?docID=20030148327.
XX
XX Sequence 14551 BP; 4362 A; 107 C; 3194 G; 6888 T; 0 U; 0 Other;
SQ
Query Match 14.7%; Score 43.2; DB 7; Length 14551;
Best Local Similarity 52.7%; Pred. No. 4.1;
Matches 117; Conservative 0; Mismatches 103; Indels 2; Gaps 1;
OY 57 TACTTCAAATATTCTTCCAGTGGTTTGTATCTTCTCAATGTTATAGTGAATATAAT 116
Db 1078 TAATACTAAACACTATATAAAAAATTTTCTTATTTAATACATCTCAAACTAAATAAAAA 1019
OY 117 AATACATGCGAAGAAACATATACTATTGTAATA--CGAAATGTGAACCTCTAT 174
Db 1018 CATACATTAATAATAAAATCAAAATAATTTAATAAAAAATTTCTAAATTTCTAAATAAAT 959
```

Qy	175	GTAAC	GCAACTCAGGCTGCGCTACCGTTT	TAATTTCTCTCTAATCTTTAA	AAACAAATA	234
Db	958	TCTAT	ATAAATCCTCTCTAGCCCAATTTCTT	CACCTAAACGTC	AAATATAAC	899
Qy	235	CTATC	TTTACTAAAGATGTTTCAGAGTAA	GATAAAATACAA	276	
Db	898	TTATT	TAATCAAAATCTTTCAAATATAAT	TAAATAAAAA	857	

Search completed: December 7, 2005, 00:02:57
Job time : 465.191 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:31:40 ; Search time 732.068 Seconds
(without alignment)
3320.997 Million cell updates/sec

Title: US-09-980-046B-2

Perfect score: 294

Sequence: 1 cggccttaaggttcctctga.....aagagagtaattccccaaaa 294

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	288.2	98.0	100445	7	US-10-322-281-170
2	47.2	16.1	15714	6	US-10-311-455-1146
3	47.2	16.1	15714	7	US-10-433-793-88
4	44.8	15.2	14924	6	US-10-311-455-197
5	44.8	15.2	14924	6	US-10-240-452-21
6	44.8	15.0	5311	6	US-10-311-455-992
7	43.6	14.8	510	6	US-10-029-386-12125
8	43.6	14.8	646	4	US-09-925-065A-467258
9	43.4	14.8	5204	6	US-10-311-455-874
10	43.4	14.8	8170	6	US-10-240-453-132
11	43.4	14.8	8423	6	US-10-311-455-1380
12	43.4	14.8	3673778	6	US-10-312-841-1
13	43.2	14.7	14551	6	US-10-240-485-137
14	43.2	14.7	3673778	6	US-10-312-841-2
15	42.6	14.5	745	5	US-10-027-632-25171
16	42.6	14.5	745	6	US-10-027-632-25171
17	42.6	14.5	5980	6	US-10-311-455-1162
18	42.6	14.5	6274	5	US-10-172-086-20
19	42.6	14.5	6274	7	US-10-221-714A-44
20	42.6	14.5	6274	7	US-10-311-507-36
21	42.6	14.5	6274	8	US-10-480-846-20
22	42.4	14.4	579	4	US-09-925-065A-842658
23	42.2	14.4	5163	6	US-10-311-455-1222

c	24	42.2	14.4	7008	7	US-10-221-714A-253	Sequence 253, App
c	25	41.8	14.2	8873	7	US-10-221-613-82	Sequence 82, Appl
c	26	41.8	14.2	11178	7	US-10-221-613-15	Sequence 15, Appl
c	27	41.6	14.1	639	7	US-10-424-599-55053	Sequence 55053, A
c	28	41.6	14.1	6104	6	US-10-311-455-1098	Sequence 1098, Ap
c	29	41.6	14.1	7829	6	US-10-311-455-1078	Sequence 1078, Ap
c	30	41.6	14.1	7829	7	US-10-221-613-150	Sequence 150, App
c	31	41.6	14.1	8056	8	US-10-473-126-386	Sequence 386, App
c	32	41.4	14.1	622	7	US-10-424-599-3380	Sequence 3380, Ap
c	33	41.4	14.1	3500	9	US-10-486-319A-75	Sequence 75, Appl
c	34	41.4	14.1	6222	6	US-10-311-455-665	Sequence 665, App
c	35	41.2	14.0	729	5	US-10-027-632-256997	Sequence 256997,
c	36	41.2	14.0	729	5	US-10-027-632-256998	Sequence 256998,
c	37	41.2	14.0	729	6	US-10-027-632-256997	Sequence 256997,
c	38	41.2	14.0	729	6	US-10-027-632-256998	Sequence 256998,
c	39	41.2	14.0	1960	3	US-09-938-842A-4687	Sequence 4687, Ap
c	40	41.2	14.0	1960	3	US-09-938-842A-4687	Sequence 4687, Ap
c	41	41.2	14.0	7456	6	US-10-311-455-1904	Sequence 1904, Ap
c	42	41.2	14.0	7456	7	US-10-240-589C-102	Sequence 102, App
c	43	41.2	14.0	8961	6	US-10-240-453-301	Sequence 301, App
c	44	41.2	14.0	18585	6	US-10-240-485-162	Sequence 162, App
c	45	41.2	14.0	118951	5	US-10-161-572-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1

US-10-322-281-170

; Sequence 170, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322.281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 170
; LENGTH: 100445
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(100445)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-170

Query Match	98.0%	Score 288.2	DB 7	Length 100445
Best Local Similarity	99.0%	Pred. No. 1.8e-56		
Matches	290	Conservative	0	Mismatches 3
			Indels	0
			Gaps	0
Qy	1	CGGCCTTAAGGTTCTCTGACAGTTGTTCTCAAGAGATCAGGATTTGAGCATCTCTTACT	60	
Db	61160	CGGCCTTAAGGTTCTCTGACAGTTGTTCTCAAGAGATCAGGATTTGAGCATCTCTTACT	61219	
Qy	61	TCAAATATTCTTCAGTGGTTTGTTCATCTTCTCAATGTTATAGTGAATAATAATA	120	
Db	61220	TCAAATATTCTTCAGTGGTTTGTTCATCTTCTCAATGTTATAGTGAATAATAATA	61279	
Qy	121	CATCAGAAAAGCATAAATCAACTTTGAATACGAAAATGTGAACCTCTATGTAAT	180	
Db	61280	CATCAGAAAAGCATAAATCAACTTTGAATACGAAAATGTGAACCTCTATGTAAT	61339	
Qy	181	GCAACTCAGGCTGCTACCGTTTTTAATTTCTCTCTATCTTTAAAAACAATATCT	240	
Db	61340	GCAACTCAGGCTGCTACCGTTTTTAATTTCTCTCTATCTTTAAAAACAATATCT	61399	
Qy	241	TTACTAAAGATGTTTTCAGAGTAAGATAAAATACAAAGAGAGTAATTCCTCAAAA	293	
Db	61400	TTACTAAAGATGTTTTCAGAGTAAGATAAAATACAAAGAGAGTAATTCCTCAAAA	61452	

```
RESULT 2
US-10-311-455-1146/c
; Sequence 1146, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Sequence of a Polynucleotide
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1146
; LENGTH: 15714
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1146

Query Match      16.1%; Score 47.2; DB 6; Length 15714;
Best Local Similarity 52.9%; Pred. No. 1.7; Indels 7; Gaps 1;
Matches 129; Conservative 0; Mismatches 108;

Qy 31 AAGAAGATCAGGATTTGAGCATCTCTTACTTCAAATTAATTTCTCCAGTGGTTTGTATC 90
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6545 ACGAAATCCAAACTTCGCTACCTTTATTTCTCTTCTTATTTCTCTCCCAATC 6486
Qy 91 TTCTCAATGTTATAGTGAATATAATATACATGCGAGAAAAAGCATAAATCAACTATTG 150
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6485 CTTTCCAACCTTAAATTTATATAAAAAAATAAAAAAATAAAAAAATAAAAAA 6426
Qy 151 AATAACGAAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCTACCGTTTAAATTT 210
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6425 AAAATCCCAAAAAAATATATATAAAAAAATAAAAAAATAAAAAAATAAAAAA 6373
Qy 211 CTCTCTAATCTTTAAAAAACAATACTATCTTTTAAAGATGTTTCAGAGTAAGATAAAA 270
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6372 TTCCCTCATATTAATTAACAATAATACCATATATACCTAAATTTAATTCACAATACATA 6313
Qy 271 TACA 274
Db |||||
6312 AAAA 6309

RESULT 3
US-10-433-793-88/c
; Sequence 88, Application US/10433793
; Publication No. US20040142334A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von mit Angiogenese assoziierten Krankheiten
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/433.793
; CURRENT FILING DATE: 2003-06-06
; NUMBER OF SEQ ID NOS: 212
; SEQ ID NO 88
; LENGTH: 15714
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-433-793-88

Query Match      16.1%; Score 47.2; DB 6; Length 15714;
Best Local Similarity 52.9%; Pred. No. 1.7; Indels 7; Gaps 1;
Matches 129; Conservative 0; Mismatches 108;

Qy 31 AAGAAGATCAGGATTTGAGCATCTCTTACTTCAAATTAATTTCTCCAGTGGTTTGTATC 90
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6545 ACGAAATCCAAACTTCGCTACCTTTATTTCTCTTCTTATTTCTCTCCCAATC 6486
Qy 91 TTCTCAATGTTATAGTGAATATAATATACATGCGAGAAAAAGCATAAATCAACTATTG 150
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6485 CTTTCCAACCTTAAATTTATATAAAAAAATAAAAAAATAAAAAAATAAAAAA 6426
Qy 151 AATAACGAAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCTACCGTTTAAATTT 210
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6425 AAAATCCCAAAAAAATATATATAAAAAAATAAAAAAATAAAAAAATAAAAAA 6373
Qy 211 CTCTCTAATCTTTAAAAAACAATACTATCTTTTAAAGATGTTTCAGAGTAAGATAAAA 270
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
6372 TTCCCTCATATTAATTAACAATAATACCATATATACCTAAATTTAATTCACAATACATA 6313
Qy 271 TACA 274
Db |||||
6312 AAAA 6309

RESULT 4
US-10-311-455-197/c
; Sequence 197, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Sequence of a Polynucleotide
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 197
; LENGTH: 14924
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: 208..209, 211, 213, 227
; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-197

Query Match      15.2%; Score 44.8; DB 6; Length 14924;
Best Local Similarity 48.1%; Pred. No. 6;
Matches 127; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

Qy 31 AAGAAGATCAGGATTTGAGCATCTCTTACTTCAAATTAATTTCTCCAGTGGTTTGTATC 90
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
12072 AATAATAAAAAATAAATTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTAA 12013
Qy 91 TTCTCAATGTTATAGTGAATATAATATACATGCGAGAAAAAGCATAAATCAACTATTG 150
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
12012 TAATCTATTTTATATCTTAAACCAATTTATATAAAAAAATAATATACAAACAAATTCAT 11953
Qy 151 AATAACGAAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCTACCGTTTAAATTT 210
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
11952 ATCCTATCAATATATATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 11893
```


Query Match 14.8%; Score 43.6; DB 6; Length 510;
Best Local Similarity 48.8%; Pred. No. 3.1;
Matches 118; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 2 GGCTTAAGGTCCTCGACAGTTGTCCTCAAGAAGATCAGGATTTGAGCATCTCTTACTT 61
DB 225 GGCATCTAAGTCCTCTCGGATGGTTCTCTTCAGTCTTACTTTATTAGATCTTCCTT 284

QY 62 CAAATTAATCTTCAGTGGTTTGGTTATCTCTCAATGTTATAGTGAATATAATATAC 131
DB 285 CAGACTCGAATTAAGTTCGTTGTTAGTCAAGATGCAATTTGATCTTTAAAAAATGGA 344

QY 122 ATGCAGAAAAAGCATAAATCAACTATTGTAATACGAAAAATGGAACCTCTATGTAACG 181
DB 345 GGAAGAAGAAGATTAATATATAGACATATTTTCAAGTTTGAAGTCCATGTAACTT 404

QY 182 CAACTCAGGTCGCTACCGTTTTTAATTTCTCTCTATCTTTTAAAAACAATATCTT 241
DB 405 TATGTTATCTGCTTTAATAACTTTTTTTTTTGAATACTTAAAAATACCCATATACTT 464

QY 242 TA 243
DB 465 CA 466

RESULT 8

US-09-925-065A-467258
; Sequence 467258, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 467258
; LENGTH: 646
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-467258

Query Match 14.8%; Score 43.6; DB 4; Length 646;
Best Local Similarity 47.4%; Pred. No. 3.4;
Matches 127; Conservative 1; Mismatches 140; Indels 0; Gaps 0;

QY 20 ACAGTTGCTCTCAAGAAGATCAGGATTTGAGCATCTCTTACTTCAAAATATTCTTCCAGTG 79
DB 251 AGACTAATGGCCAAAGTACATTTTATTAATGCGTCTTACTAATGATTTTAAATGA 310

QY 80 GTTTTGGTTATCTTCTCAATGTTATAGTGAATATAATAATACATGCAGAAAAAGCATAA 139
DB 311 ATAAATCAATTTAAAAAATGAGCAAGATGATTGATAATTCACACAAGAAGAAATACAA 370

QY 140 ATCAACTATTGAATAACGAAATGTGAACCTCTATGTAAGTCAACTCAGCGTCCCTACC 199
DB 371 AYGAAACCAAGACACTTGTGTCCTATTAGTGAATATGAATGCAAGTAAAAACAACAATA 430

QY 200 GTTTTAAATTTCTCTCTAAATCTTTAAAAACAATACTATCTTTACTAAAGATGTTTCAGA 259

DB 431 AAATATAACTTTTACACAAATTTTCAAAAGTATTGGTAATATATATTGTTGGTG 490

QY 260 GTAAGTATAAATACAAAGAGAGCTAATTC 287

DB 491 GGAATTTTAAATTTTAAAAATAGGAACCTC 518

RESULT 9

US-10-311-455-874/c
; Sequence 874, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 874
; LENGTH: 5204
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-874

Query Match 14.8%; Score 43.4; DB 6; Length 5204;
Best Local Similarity 50.2%; Pred. No. 8.5;
Matches 107; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 82 TTGTTTATCTTCTCAATGTTATAGTGAATATAATATCATGCAGAAAAAGCATAAAT 141
DB 2290 TTTTTCACGTAACGGCCTAATAACAAAAAATAAAAAAATAAAAAAATAAAAAA 2231

QY 142 CAACTATTGAATAACGAAATGTGAACCTCTGTAACTGCACTCAGCTGCTCCGT 201
DB 2230 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 2171

QY 202 TTTTAAATTTCTCTCTAATCTTTAAAAACAATACTATCTTTACTAAAGATGTTTCAGAGT 261
DB 2170 TAAATTCATTTCTTAAATCTTATCTATCTATCATATATAATTTTAAAAAACGTTCTCAAACT 2111

QY 262 AGATAAATAACAAAGAGAGTAATCCCAAAA 294
DB 2110 AAAAAAATAAATACTATAATAATAACCAAAATA 2078

RESULT 10

US-10-240-453-132/c
; Sequence 132, Application US/10240453
; Publication No. US20030148326A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription
; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated with DNA Transcription
; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240,453
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:34:11 ; Search time 200.691 Seconds
(without alignments)
455.919 Million cell updates/sec

Title: US-09-980-046B-2
Perfect score: 294
Sequence: 1 cggccttaaggcttcctctga.....aagagagtaattcccaaaaa 294

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA_New:*
1: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB_seq.*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB_seq.*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB_seq.*
4: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB_seq.*
5: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB_seq.*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB_seq.*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB_seq.*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB_seq.*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB_seq.*
10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB_seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	39.4	13.4	4497	6	US-10-518-753-15
C 2	36.8	12.5	2467	6	US-10-750-185-29248
C 3	36	12.2	1857	6	US-10-750-185-42217
C 4	35.6	12.1	1740	6	US-10-750-185-37198
C 5	35.6	12.1	2536	7	US-11-087-227-9
C 6	35.6	12.1	2613	7	US-11-087-227-7
C 7	35.6	12.1	2748	7	US-11-087-227-5
C 8	35.4	12.0	1899	6	US-10-750-185-62020
C 9	35.2	12.0	2444	6	US-10-750-185-29713
C 10	35.2	12.0	5501	6	US-10-750-185-28699
C 11	35	11.9	3069	6	US-10-750-185-4352
C 12	35	11.9	3206	6	US-10-750-185-4382
C 13	34.8	11.8	1352	7	US-11-096-622-17
C 14	34.8	11.8	1757	7	US-11-096-622-14
C 15	34.8	11.8	8512	7	US-11-096-622-20
C 16	34.8	11.8	8565	7	US-11-096-622-21
C 17	34.6	11.8	4731	6	US-10-750-185-60556
C 18	34.2	11.6	2253	6	US-10-750-185-60556
C 19	34.2	11.6	2311	6	US-10-485-517-68
C 20	34.2	11.6	3416	6	US-10-750-185-3946
C 21	34.2	11.6	3520	6	US-10-750-185-3657
C 22	34	11.6	906	6	US-10-750-185-205
C 23	34	11.6	2625	6	US-10-750-185-3777

C 24	34	11.6	3383	6	US-10-750-185-64159	Sequence 64159, A
C 25	33.8	11.5	2017	6	US-10-750-185-37628	Sequence 37628, A
C 26	33.8	11.5	3376	6	US-10-750-185-3445	Sequence 3445, A
C 27	33.6	11.4	1045	6	US-10-750-185-27306	Sequence 27306, A
C 28	33.6	11.4	1124	6	US-10-750-185-27308	Sequence 27308, A
C 29	33.6	11.4	1837	6	US-10-750-185-53684	Sequence 53684, A
C 30	33.2	11.3	938	6	US-10-750-185-60864	Sequence 60864, A
C 31	33.2	11.3	2318	6	US-10-750-185-44654	Sequence 44654, A
C 32	33.2	11.3	4597	6	US-10-518-753-4	Sequence 4, Appli
C 33	33.2	11.3	6220	6	US-10-518-753-3	Sequence 3, Appli
C 34	33	11.2	14911	7	US-11-112-908-63	Sequence 63, Appl
C 35	32.8	11.2	1305	6	US-10-750-185-48543	Sequence 48543, A
C 36	32.8	11.2	1507	6	US-10-750-185-30562	Sequence 30562, A
C 37	32.8	11.2	1526	6	US-10-750-185-36498	Sequence 36498, A
C 38	32.8	11.2	2316	6	US-10-750-185-27930	Sequence 27930, A
C 39	32.8	11.2	3183	6	US-10-750-185-4416	Sequence 4416, Ap
C 40	32.8	11.2	3546	6	US-10-750-185-3457	Sequence 3457, Ap
C 41	32.8	11.2	3852	6	US-10-750-185-3374	Sequence 3374, Ap
C 42	32.8	11.2	3932	6	US-10-750-185-3420	Sequence 3420, Ap
C 43	32.8	11.2	159497	7	US-11-112-908-61	Sequence 61, Appl
C 44	32.6	11.1	1244	6	US-10-750-185-24614	Sequence 24614, A
C 45	32.6	11.1	1385	6	US-10-750-185-47233	Sequence 47233, A

ALIGNMENTS

RESULT 1

US-10-518-753-15/c
; Sequence 15, Application US/10518753
; Publication No. US20050262589A1
; GENERAL INFORMATION:
; APPLICANT: Calgene LLC
; TITLE OF INVENTION: Nucleic Acid Sequences and Methods of Use for the Production of P
; FILE REFERENCE: 16518.129
; CURRENT APPLICATION NUMBER: US/10/518.753
; CURRENT FILING DATE: 2004-12-21
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 4497
; TYPE: DNA
; ORGANISM: Glycine max
US-10-518-753-15

Query Match 13.4%; Score 39.4; DB 6; Length 4497;
Best Local Similarity 55.4%; Pred. No. 1.5;
Matches 97; Conservative 0; Mismatches 76; Indels 2; Gaps 1;

Qy	106	TGAATAATAATACATGCAGAAAGCAATCACTATTGAATAACGA--AAATG	163
Db	3336	TGATATATTCAAGACATATATTATAAAATTAAGTTTAAATTAGTAAATTTAAATTT	3277
Qy	164	TGAACCTCTATGTAACTGCAACTCAGCGTCCCTACCGTTTTTAATTTCTCTCTAATCTTT	223
Db	3276	TGAATCTAAAATCAGTCCAAATTTAGTGTGCTATATTTTAATTTTCAATAGCAATTT	3217
Qy	224	AAAAACAAATCTATCTTTTACTAAAGATGTTTCAGAGTAAGATAAATACAAAGA	278
Db	3216	TATTACAAATTTTCAGTTTAAATGAATCAATTTGATTTTAATTAACCTTGTAAGA	3162

RESULT 2

US-10-750-185-29248
; Sequence 29248, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom

```
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29248
; LENGTH: 2467
; TYPE: DNA
; ORGANISM: Bovine 19866881228615
US-10-750-185-29248

Query Match      12.5%; Score 36.8; DB 6; Length 2467;
Best Local Similarity 50.0%; Pred. No. 5.2;
Matches 92; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 101 TATAGTGAATATAATATACATGCAGAAAAAGACATAAATCAACTATTGGAATACGAAA 160
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
505 TATTGTAAGTTAAAAATAAATAATTAATAAATAATAGAAAAATCAAAAAATATATATAT 564
Qy 161 ATGTGAACCTCTATGTAACTGCAACTCAGCGTGCCTACCGTTTTTAATTTCTCTAATC 220
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
565 ATTAATACCTGATTTAATTCATCTTAATAAACTACTGGTGTGGTTTAAATTTTTT 624
Qy 221 TTTAAAAACAATACTACTTTTACTAAAGATGTTTTCAGAGTAAGATAAAAAACAAAAGAGA 280
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
625 TTTTAATTAATTAATTTAATTTGGAGGTTTTTAATTAATAAAAAAAGAGAGCAGA 684
Qy 281 GTAA 284
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
685 ATAA 688

RESULT 3
US-10-750-185-42217/C
; Sequence 42217, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42217
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Bovine 19866880837875
US-10-750-185-42217

Query Match      12.2%; Score 36; DB 6; Length 1857;
Best Local Similarity 55.6%; Pred. No. 7.3;
Matches 90; Conservative 0; Mismatches 70; Indels 2; Gaps 1;

Qy 6 TTAGGTTCTCTGACAGTTGTCTCAAGAAGATCAGGATTTGAGCATCTCTACTTCAAA 65
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1766 TTACGTCTCTCTTCACTTTCATCAAGAGCGCTTTTAGTTCCTCTTCACTTTCTGCCA- 1708
Qy 66 TTATTTCTCCAGTGGTTTGTGTATCTTCTCAATGTTATAGTGAATATAATAATACATGC 125
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 1707 -TAAGAGTGGTGTCTATCTGCATATCTGTGTAGTACAGAGAGATATAATAATGAATTG 1649
Qy 126 AGAAAAAGCATAAATCAACTATTGTAATACGAAATGTGAA 167
Db 1648 TGGCAAAAAATGTAATAATCTGCTGAATCTGGGGAAGAGAA 1607

RESULT 4
US-10-750-185-37198/C
; Sequence 37198, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37198
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Bovine 19866881201580
US-10-750-185-37198

Query Match      12.1%; Score 35.6; DB 6; Length 1740;
Best Local Similarity 49.0%; Pred. No. 8.9;
Matches 95; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy 82 TTTGTTATCTTCTCAATGTTATAGTGAAATATAATAATATATATATATATATATATAT 141
Db 284 TTTGTTCTATTGTTATGTTTCCAGGCAATAAATAAATCTGGTGACATATCATGGCTATT 225
Qy 142 CAACTATTGAATAACGAAAAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCTACCGT 201
Db 224 TAAATTTATATAGTGTAAATTTTAAATCTTTTAAAGTAGTATTTCATGTATAGTACTGT 165
Qy 202 TTTTAATTTCTCTCTAATCTTTTAAACCAAAATCTATCTTTACTAAGATGTTTCAGAGT 261
Db 164 ACTAAATTTAAGATTATTGCTTATTTTCCACCAAAAAATCTAACACAGTAGGTTCTGGGCAT 105
Qy 262 AAGATAAAATACAA 275
Db 104 TATTTAAATATAAA 91

RESULT 5
US-11-087-227-9
; Sequence 9, Application US/11087227
; Publication No. US20050260566A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Timothy J.
; APPLICANT: Malinowski, Douglas P.
; APPLICANT: Taylor, Adriann J.
; APPLICANT: Parker, Margaret R.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; DETECTION OF CERVICAL DISEASE
; FILE REFERENCE: 046143/287139
; CURRENT APPLICATION NUMBER: US/11/087,227
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: 60/556,495
; PRIOR FILING DATE: 2004-03-24
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
```

```
; LENGTH: 2536
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-087-227-9

Query Match      12.1%; Score 35.6; DB 7; Length 2536;
Best Local Similarity 55.7%; Pred. No. 9.7;
Matches 68; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 159 AAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCTACCGTTTAAATTTCTCTCTAA 218
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2415 AATTGTGACATATGTATTATATAAACTTAACCTTTTAACTACTGTTTATTTTAGCCCAT 2474

Qy 219 TCTTTAAAAACAATACTATCTTTACTAAAGATGTTTCAGAGTAAGATAAAATACAAGA 278
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2475 TGTTTAAAAAATAAAGTTAAAAAAATTTAACTGCTTAAAGTAAGAAAAAATAAAAAA 2534

Qy 279 GA 280
Db 2535 AA 2536

RESULT 6
US-11-087-227-7
; Sequence 7, Application US/11087227
; Publication No. US20050260566A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Timothy J.
; APPLICANT: Malinowski, Douglas P.
; APPLICANT: Taylor, Adriann J.
; APPLICANT: Parker, Margaret R.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: DETECTION OF CERVICAL DISEASE
; FILE REFERENCE: 046143/287139
; CURRENT APPLICATION NUMBER: US/11/087,227
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: 60/556,495
; PRIOR FILING DATE: 2004-03-24
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2613
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-087-227-7

Query Match      12.1%; Score 35.6; DB 7; Length 2613;
Best Local Similarity 55.7%; Pred. No. 9.8;
Matches 68; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 159 AAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCTACCGTTTAAATTTCTCTCTAA 218
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2492 AATTGTGACATATGTATTATATAAACTTAACCTTTTAACTACTGTTTATTTTAGCCCAT 2551

Qy 219 TCTTTAAAAACAATACTATCTTTACTAAAGATGTTTCAGAGTAAGATAAAATACAAGA 278
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2552 TGTTTAAAAAATAAAGTTAAAAAAATTTAACTGCTTAAAGTAAGAAAAAATAAAAAA 2611

Qy 279 GA 280
Db 2612 AA 2613

RESULT 7
US-11-087-227-5
; Sequence 5, Application US/11087227
; Publication No. US20050260566A1
; GENERAL INFORMATION:
; APPLICANT: Fischer, Timothy J.
; APPLICANT: Malinowski, Douglas P.
; APPLICANT: Taylor, Adriann J.
; APPLICANT: Parker, Margaret R.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
```

```
; TITLE OF INVENTION: DETECTION OF CERVICAL DISEASE
; FILE REFERENCE: 046143/287139
; CURRENT APPLICATION NUMBER: US/11/087,227
; CURRENT FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: 60/556,495
; PRIOR FILING DATE: 2004-03-24
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2748
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-087-227-5

Query Match      12.1%; Score 35.6; DB 7; Length 2748;
Best Local Similarity 55.7%; Pred. No. 9.9;
Matches 68; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 159 AAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCTACCGTTTAAATTTCTCTCTAA 218
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2627 AATTGTGACATATGTATTATATAAACTTAACCTTTTAACTACTGTTTATTTTAGCCCAT 2686

Qy 219 TCTTTAAAAACAATACTATCTTTACTAAAGATGTTTCAGAGTAAGATAAAATACAAGA 278
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2687 TGTTTAAAAAATAAAGTTAAAAAAATTTAACTGCTTAAAGTAAGAAAAAATAAAAAA 2746

Qy 279 GA 280
Db 2747 AA 2748

RESULT 8
US-10-750-185-62020
; Sequence 62020, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: WMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62020
; LENGTH: 1899
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-62020

Query Match      12.0%; Score 35.4; DB 6; Length 1899;
Best Local Similarity 54.1%; Pred. No. 10;
Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Qy 113 TAATAATACATGCAGAAAAAGCATAAATCAACTATTGAAATACGAAAAATGTGAACCTCT 172
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1079 TAATGATAAGGGAAAAAAATTAATAAAAAAAGATGATAAAGGAAGCTTCAAAACCT 1138

Qy 173 ATGTAACCTGCAACTCAGGCTGCCTACCGTTTAAATTTCTCTCTAAATTTTAAACAAA 232
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1139 ATGTACATCTCTTCAAGAAATCCCTTCTAGATTTTGTGTATCTGTCTACTTTCACACACTA 1198

Qy 233 TACTATCTTTACT 245
   ||||| ||||| |||||
Db 1199 AACATATAACCT 1211
```

```
RESULT 9
US-10-750-185-29713/c
; Sequence 29713, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29713
; LENGTH: 2444
; TYPE: DNA
; ORGANISM: Bovine 19866880767225
US-10-750-185-29713

Query Match      12.0%; Score 35.2; DB 6; Length 2444;
Best Local Similarity 48.5%; Pred. No. 12;
Matches 97; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 34 AAGATCAGGATTTGAGATCTCTTACTTCAAAATTATCTTCAGTGGTTTGTATCTTC 93
DB 2195 AAGCCAAAGATCCAAGCTCTCTCTTAACTTTTAGATTGATGGTTCTCTGTGCC 2136

QY 94 TCAATGTTATAGTGAATATAAATACATGCGAGAAAAGACATAAATCACTATTGAAT 153
DB 2135 ACAATATAAATCATGCTTCAGGATCTTCATCTTGGGAAAGAAAACAGAAAGAACAT 2076

QY 154 AAGCAAAATGTGAACCTCTATGTAACCTGCAACTCAGGCTGCCCTACCGTTTTTAATTCTC 213
DB 2075 CCATCCAGGTCTCCCTTTATCTGTTGCCCTTCTCGCATGCACCTGCTCTCTA 2016

QY 214 TCTAATCTTTAAAACAAAT 233
DB 2015 TTTCTTCTCAAAAGAAAAAT 1996

RESULT 10
US-10-750-185-28699
; Sequence 28699, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28699
; LENGTH: 5501
; TYPE: DNA
; ORGANISM: Bovine 19866880637762
US-10-750-185-28699

Query Match      12.0%; Score 35.2; DB 6; Length 5501;
Best Local Similarity 52.8%; Pred. No. 15;
Matches 76; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 137 TAAATCAACTATTGAAATAACGAAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCT 196
DB 4128 TAGATAACTAATGAACAACACAGACAAAGACTCTTATGGACTTTAAATGCACACAGAAAT 4187

QY 197 ACCGTTTTTAATTTCTCTCTAATCTTTAAAAACAAATATCTATCTTTACTAAAGATGTTTC 256
DB 4188 GTATATTTAAAAATCCTAAACTATACATTAATTTAGCAATACATATATTTATCATTTCTC 4247

QY 257 AGAGTAAAGATAAAATACAAAGAGA 280
DB 4248 AATACAGAAAAAAGACATAATGA 4271

RESULT 11
US-10-793-626-4352
; Sequence 4352, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4352
; LENGTH: 3069
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4352

Query Match      11.9%; Score 35; DB 6; Length 3069;
Best Local Similarity 55.3%; Pred. No. 14;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 65 ATTATTCTTCCAGTGGTTTGTGTATCTTCTCAATGTTATAGTGAATATAATATATCATG 124
DB 2418 ACTGTCACCTGTGCTTTTATTAATAATTTATTAGTAAATAAATAATAAATTTTACAG 2477

QY 125 CAGAAAAAGCATAAATCAACTATTGAAATACGAAAAATGGAACCTCTATGTAACGCA 184
DB 2478 AAAAAAATATCAAAAAACCAAAATATAAAAAATGAAATGATGGATTAGGATGCTTTTGCTA 2537

QY 185 CTC 187
DB 2538 CAC 2540

RESULT 12
US-10-793-626-4382/c
; Sequence 4382, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4382
```

```
; LENGTH: 3206
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; US-10-793-626-4382

Query Match      11.9%; Score 35; DB 6; Length 3206;
Best Local Similarity 55.3%; Pred. No. 14;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 65 ATTATCTTCAGTGGTTTCTTCAATGTTATAGTGAATAATAATACATG 124
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1064 ACTGCCACCTTGCTTTATTAAAAATTATTAGTAAATATAAAATTTACAG 1005
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 125 CAGAAAACGATTAATCACTATTGATACGAAATGCAACCTCTATGTAACGAA 184
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1004 AAAAAAATATCAAAACCAAAATATAAAATGAAATGATGGATGCTTTTGCTA 945
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 185 CTC 187
    ||
Db 944 CAC 942
    ||

RESULT 13
US-11-096-622-17/c
; Sequence 17, Application US/11096622
; Publication No. US20050260643A1
; GENERAL INFORMATION:
; APPLICANT: HUNG, MIEN-CHIE
; APPLICANT: DAY, CHI-PING
; APPLICANT: RAU, KUN-MING
; APPLICANT: XIE, XIAOMING
; APPLICANT: LI, ZHENG
; TITLE OF INVENTION: CANCER SPECIFIC PROMOTERS
; FILE REFERENCE: UTSC:845US
; CURRENT APPLICATION NUMBER: US/11/096,622
; PRIOR FILING DATE: 2005-04-01
; PRIOR APPLICATION NUMBER: 60/559,111
; PRIOR FILING DATE: 2004-04-02
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 1352
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-11-096-622-17

Query Match      11.8%; Score 34.8; DB 7; Length 1352;
Best Local Similarity 50.6%; Pred. No. 13;
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 78 TGGTTTCTTATCTTCAATGTTATAGTGAATAATAATACATGCAAAAAAGCAT 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1267 TTGTTTCGTTTGTGTTGATGAGCGTATGTTAGTACTATCGATTACACAAAAACC 1208
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 138 AAATCAACTATTGAATAACGAAATGTGAACCTCTATGTAACCTCAGGCTGCCTA 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1207 AACACACAGATGTAATGAATAAAGATATTTTATGCGCGCGCTCTAGAAATCTA 1148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 198 CGGTTTTTAATTTCTCTCTAATCTTTTAAACAAATACTATCTTTA 243
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1147 GCTGATATAGTGTGGCTCAAAACCTTCAGCACAATCACACCGTTA 1102
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
US-11-096-622-14/c
; Sequence 14, Application US/11096622
; Publication No. US20050260643A1
```

```
; GENERAL INFORMATION:
; APPLICANT: HUNG, MIEN-CHIE
; APPLICANT: DAY, CHI-PING
; APPLICANT: RAU, KUN-MING
; APPLICANT: XIE, XIAOMING
; APPLICANT: LI, ZHENG
; TITLE OF INVENTION: CANCER SPECIFIC PROMOTERS
; FILE REFERENCE: UTSC:845US
; CURRENT APPLICATION NUMBER: US/11/096,622
; PRIOR FILING DATE: 2005-04-01
; PRIOR APPLICATION NUMBER: 60/559,111
; PRIOR FILING DATE: 2004-04-02
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1757
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-11-096-622-14

Query Match      11.8%; Score 34.8; DB 7; Length 1757;
Best Local Similarity 50.6%; Pred. No. 14;
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 78 TGGTTTCTTATCTTCAATGTTATAGTGAATAATAATACATGCAAAAAAGCAT 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1672 TTGTTTCGTTTGTGTTGATGAGCGTATGTTAGTACTATCGATTACACAAAAACC 1613
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 138 AAATCAACTATTGAATAACGAAATGTGAACCTCTATGTAACCTCAGGCTGCCTA 197
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1612 AACACACAGATGTAATGAATAAAGATATTTTATGCGCGCGCTCTAGAAATCTA 1553
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 198 CGGTTTTTAATTTCTCTCTAATCTTTTAAACAAATACTATCTTTA 243
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1552 GCTGATATAGTGTGGCTCAAAACCTTCAGCACAATCACACCGTTA 1507
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
US-11-096-622-20/c
; Sequence 20, Application US/11096622
; Publication No. US20050260643A1
; GENERAL INFORMATION:
; APPLICANT: HUNG, MIEN-CHIE
; APPLICANT: DAY, CHI-PING
; APPLICANT: RAU, KUN-MING
; APPLICANT: XIE, XIAOMING
; APPLICANT: LI, ZHENG
; TITLE OF INVENTION: CANCER SPECIFIC PROMOTERS
; FILE REFERENCE: UTSC:845US
; CURRENT APPLICATION NUMBER: US/11/096,622
; CURRENT FILING DATE: 2005-04-01
; PRIOR APPLICATION NUMBER: 60/559,111
; PRIOR FILING DATE: 2004-04-02
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 8512
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; US-11-096-622-20

Query Match      11.8%; Score 34.8; DB 7; Length 8512;
Best Local Similarity 50.6%; Pred. No. 20;
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 78 TGGTTTCTTATCTTCAATGTTATAGTGAATAATAATACATGCAAAAAAGCAT 137
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 2952 TTGTTTCGTTTTGTTTTGATGGAGCGGTATGTTAGTACTATCGATTTCACACAAAAACC 2893
Qy 138 AAATCAACTATTGAATAACGAAATGTGAACCTCTATGTAACTGCAACTCAGGCTGCCTA 197
Db 2892 AACACACAGATGTAATGAAATAAAGATATTTTATTGCGGCCGCCGCTCTAGAAATCTA 2833
Qy 198 CCGTTTTTAATTTCTCTCTAATCTTTAAAAACAAATACTATCTTTA 243
Db 2832 GCTGATATAGTGTGGCTCAAAACCTTCAGCACAANTCACCCTTA 2787

Search completed: December 6, 2005, 23:16:50
Job time : 201.691 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 124.194 Seconds
(without alignments)
4239.411 Million cell updates/sec

Title: US-09-980-046b-3

Perfect score: 79

Sequence: 1 cgggaatcccccttcagttc.....aaacaacgactctaaaaa 79

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001as.*
- 5: Geneseq2001bs.*
- 6: Geneseq2002as.*
- 7: Geneseq2002bs.*
- 8: Geneseq2003as.*
- 9: Geneseq2003bs.*
- 10: Geneseq2003cs.*
- 11: Geneseq2003ds.*
- 12: Geneseq2004as.*
- 13: Geneseq2004bs.*
- 14: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	100.0	79	4	AAC89734
2	75.8	95.9	1464	3	AAC59772 Human sec
3	74.8	94.7	428	13	ADU14748
4	74.8	94.7	522	14	AEA36210
5	74.8	94.7	592	10	ADF81837
6	74.8	94.7	1985	12	ADQ86642
7	74.8	94.7	1985	13	ADQ85525
8	74.8	94.7	2000	4	AAC88152
9	74.8	94.7	2000	6	ABL65927
10	74.8	94.7	2000	6	ABK84431
11	74.8	94.7	2000	12	ADJ74845
12	74.8	94.7	2000	13	ADP52746
13	74.8	94.7	2000	13	ADP24663
14	74.8	94.7	2000	13	ADU06261
15	74.8	94.7	2000	14	ADW38555
16	74.8	94.7	2180	2	AAV32446
17	71	89.9	90	4	AAC89735
18	63.4	80.3	212	2	AAT26356
19	54	68.4	401	10	ADF80286

20	29.6	37.5	110000	6	ABA03041_08	Continuation (9 of
21	29.2	37.0	2278	6	ABQ70619	Abq70619 Listeria
22	29	36.7	528	5	ADL43615	Adl43615 Human ova
23	28.8	36.5	16037	4	AAK78529	AAk78529 Human imm
24	28.8	36.5	16037	6	ABK69947	ABk69947 Human sec
25	28.8	36.5	16037	8	ABZ74623	Abz74623 Secreted
26	28.8	36.5	16037	10	ADC21014	Adc21014 Human sec
27	28.8	36.5	16037	10	ABZ68144	Abz68144 Human sec
28	28.8	36.5	16107	4	AAK78528	AAk78528 Human imm
29	28.8	36.5	16107	6	ABK69948	ABk69948 Human sec
30	28.8	36.5	16107	8	ABZ74624	Abz74624 Secreted
31	28.8	36.5	16107	10	ADC21015	Adc21015 Human sec
32	28.8	36.5	16107	10	ABZ68145	Abz68145 Human sec
33	28.8	36.5	31828	11	ACN44338	ACn44338 Human gen
34	28.4	35.9	110000	14	ADZ13631_2	Continuation (3 of
35	28.4	35.9	110000	14	ADZ13631_3	Continuation (4 of
36	28.4	35.9	110000	14	ADZ13620_2	Continuation (3 of
37	28.4	35.9	110000	14	ADZ13620_3	Continuation (4 of
38	28.4	35.9	299598	12	ADQ59380	Adq59380 Human can
39	28	35.4	683	5	ADL37231	Adl37231 Human ova
40	28	35.4	683	5	ADI72082	Adi72082 Human ova
41	28	35.4	4722	14	ADW10342	Adw10342 Colon pro
42	28	35.4	17144	4	AAS46665	Aas46665 Tumour su
43	27.8	35.2	495	4	ABA55327	Abas55327 Human foe
44	27.8	35.2	495	4	AAI34978	AAi34978 Probe #36
45	27.8	35.2	495	4	ABA44865	Abas44865 Human bre

ALIGNMENTS

RESULT 1

AAC89734

ID AAC89734 standard; cDNA; 79 BP.

XX

AC AAC89734;

XX

12-MAR-2001 (first entry)

XX

Human gastrointestinal inflammation-related CDNA, SEQ ID NO: 3.

XX

Human; cytostatic; immunomodulator; immunostimulant; vulnery;

KW anti-inflammatory; neuroprotective; antibacterial; gene therapy;

KW gastrointestinal inflammation; immune system disorder; genetic disorder;

KW cancer; autoimmune disorder; infection; wound healing; ss.

XX

OS Homo sapiens.

XX

WO200073324-A2.

XX

07-DEC-2000.

XX

01-JUN-2000; 2000WO-US015191.

XX

01-JUN-1999; 99US-0137058P.

XX

(DIGI-) DIGITAL GENE TECHNOLOGIES INC.

XX

Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Hasel KW;

XX

WPI; 2001-061508/07.

XX

New polynucleotides and polypeptides, useful in gene therapy and in diagnosing a pathological condition, e.g. for modulating gene expression in disorders.

PT

Claim 1; Page 84; 108pp; English.

XX

The present sequence is one of a number of isolated human polynucleotides which are useful in gene therapy, and for diagnosing a pathological condition or a susceptibility to it. In particular, the polynucleotides are useful for modulating gene expression in gastrointestinal

CC inflammation. The polynucleotides are useful for chromosome
 CC identification, controlling gene expression through triple helix
 CC formation or antisense DNA or RNA, or identifying individuals from minute
 CC biological samples using DNA-based identification techniques. The
 CC polynucleotides can also be used as an alternative to restriction
 CC fragment length polymorphism (RFLP), by determining the actual base-by-
 CC base DNA sequences of selected portions of an individual's genome. The
 CC polynucleotides may also be used as molecular weight markers on Southern
 CC gels, as diagnostic probes for the presence of a specific mRNA, as a
 CC probe to subtract-out known sequences in the process of discovering novel
 CC polynucleotides, or as an antigen to elicit an immune response. The
 CC polypeptides are useful in diagnostic procedures to detect a disorder.
 CC The polynucleotides and polypeptides are useful for preventing, treating
 CC or ameliorating immune system disorders, genetic disorders, cancers, some
 CC autoimmune disorders, or infections. The polynucleotides and polypeptides
 CC are also useful for differentiating, proliferating or attracting cells,
 CC leading to the regeneration of tissues, especially in wounds or burns.
 CC The polypeptides and polynucleotides may also be used as a food additive
 CC or preservative
 XX
 SQ Sequence 79 BP; 31 A; 18 C; 12 G; 18 T; 0 U; 0 Other;

Query Match 100.0%; Score 79; DB 4; Length 79;
 Best Local Similarity 100.0%; Pred. NO. 1.8e-15;
 Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 DB 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 QY 61 AACAAACCGACTCTAAAAA 79
 DB 61 AACAAACCGACTCTAAAAA 79

RESULT 2
 AAC59772
 ID AAC59772 standard; cDNA; 1464 BP.

AC AAC59772;

XX 26-JAN-2001 (first entry)

XX Human secreted protein gene 35 SEQ ID NO:45.

XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
 KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
 KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
 KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
 KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;
 KW cerebrovascular disorder; angiogenesis; nervous system disorder;
 KW ocular disorder; wound healing; skin aging; food additive; preservative;
 KW ss.

XX Homo sapiens.

XX WO200056751-A1.

XX 28-SEP-2000.

XX 09-MAR-2000; 2000WO-US006013.

XX 19-MAR-1999; 99US-0125360P.

XX 11-JUN-1999; 99US-0138626P.

XX 03-DEC-1999; 99US-0168662P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Reen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-579482/54.

XX P-PSDB; AAB34611.

PT Isolated nucleic acid molecule encoding a human secreted protein is used
 PT in preventing, treating or ameliorating a medical condition.
 XX
 PS Claim 1; Page 362; 419pp; English.

XX The polynucleotide sequences given in AAC59738 to AAC59787 encode the
 CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to
 CC AAB34686 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Example of activities include:
 CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;
 CC cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
 CC neuroprotective; antibacterial; virucide; fungicide; and
 CC ophthalmological. The polynucleotides and proteins can be used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 CC in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and
 CC cancers of the breast or liver, cardiovascular disorders, cerebrovascular
 CC disorders, angiogenesis, nervous system disorders, infections caused by
 CC bacteria, viruses and fungi and ocular disorders. The proteins can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The proteins can also be used as a
 CC food additive or preservative to increase or decrease storage
 CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used
 CC in the exemplification of the present invention

XX Sequence 1464 BP; 309 A; 483 C; 408 G; 264 T; 0 U; 0 Other;

Query Match 95.9%; Score 75.8; DB 3; Length 1464;
 Best Local Similarity 97.5%; Pred. NO. 3.4e-14;
 Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 DB 1071 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1130
 QY 61 AACAAACCGACTCTAAAAA 79
 DB 1131 AACAAACCGACTCTAAAAA 1149

RESULT 3

ADU14748

ID ADU14748 standard; DNA; 428 BP.

XX ADU14748;

XX 27-JAN-2005 (first entry)

XX Solid tumour prognosis gene seqid 5187.

XX cytostatic; gene therapy; expression profile; solid tumour;

KW peripheral blood mononuclear cell; PBMC; prognosis; ds.

XX Unidentified.

XX WO2004097052-A2.

XX 11-NOV-2004.

XX 29-APR-2004; 2004WO-US013587.

XX 29-APR-2003; 2003US-0466067P.

XX 23-JAN-2004; 2004US-0538246P.

XX (AMHP) WYETH.

XX (STRA/) STRAHS A.

PI Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
 PI Immerman F, Dorner AJ;
 XX WPI; 2004-804779/79.
 DR
 XX
 XX A method, useful for prognosing and treating solid tumor, comprises
 PT comparing an expression profile of a gene expressed in peripheral blood
 PT mononuclear cells to a reference expression profile of a gene.
 XX
 XX Disclosure; Page; 111pp; English.
 PS
 XX
 CC The invention describes a method comprising comparing an expression
 CC profile of at least one gene in a peripheral blood sample of a patient to
 CC at least one reference expression profile of the at least one gene, where
 CC the patient has a solid tumor, and each of the gene is differentially
 CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
 CC of patients as compared to PBMCs of a second class of patients, where
 CC both the first and second classes of patients have the solid tumor, and
 CC each of the first and second classes is a subcluster formed by an
 CC unsupervised clustering analysis of gene expression profiles in PBMCs of
 CC a population of patients who have the solid tumor, and where the
 CC majority of the first class of patients has a first clinical outcome, and
 CC the majority of the second class of patients has a second clinical
 CC outcome. Also described are: a system comprising (i) a memory or a
 CC storage medium including data that represent an expression profile of at
 CC least one gene in a peripheral blood sample of a patient who has a solid
 CC tumor, (ii) at least another storage medium including data that
 CC represent at least one reference expression profile of the gene, (iii) a
 CC program capable of comparing the expression profile to the reference
 CC expression profile, and (iv) a processor capable of executing the
 CC program, where expression levels of the gene in peripheral blood
 CC mononuclear cells of patients who have the solid tumor correlate with
 CC clinical outcomes of the patients; and a nucleic acid or protein array
 CC comprising concentrated probes for solid tumor prognosis genes, where
 CC each of the solid tumor prognosis genes is differentially expressed in
 CC PBMCs of a first class of patients as compared to PBMCs of a second class
 CC of patients, where both the first and second classes of patients have a
 CC solid tumor, and where the first class of patients has a first clinical
 CC outcome, and the second class of patients has a second clinical outcome.
 CC The method, system, and array are useful for prognosing and treating
 CC solid tumors. This sequence represents a solid tumor prognosis gene of
 CC the invention. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/international_published_pat_sequences.
 XX
 SQ Sequence 428 BP; 107 A; 108 C; 112 G; 81 T; 0 U; 20 Other;
 Query Match 94.7%; Score 74.8; DB 13; Length 428;
 Best Local Similarity 97.4%; Pred. No. 5.5e-14;
 Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CGGGATCCCTTCAGTCTTTGAAAAGTTCATGACTCAATATCTGAATGAAGAA 60
 Db 100 CGGGATCCCTTCAGTCTTTGAAAAGTTCATGACTCAATATCTGAATGAAGAA 159
 QY 61 AACAAACCGACTCTAAA 78
 Db 160 AACAAACCGACTCAAAA 177
 RESULT 4
 ID AEA36210/c
 XX AEA36210 standard; DNA; 522 BP.
 AC AEA36210;
 XX
 DT 25-AUG-2005 (first entry)
 XX
 DE Human nucleic acid sequence #142.
 XX
 KW Screening; gene expression; colorectal tumor; colitis; Crohn's disease;
 KW irritable bowel syndrome; gastrointestinal disease; cytostatic;
 KW gastrointestinal-gen.; antiinflammatory; ds.

XX Homo sapiens.
 OS WO2005054507-A2.
 PN 16-JUN-2005.
 PD 03-DEC-2004; 2004WO-GB005078.
 PF 04-DEC-2003; 2003GB-00028048.
 PR (UYSH-) UNIV SHEFFIELD.
 PA Corfe B, Chirakkal H;
 PI WPI; 2005-435407/44.
 DR
 XX Screening for nucleic acid molecules exhibiting altered expression in
 PT cells grown in the presence of butyrate, and detection of the nucleic
 PT acid molecules or the encoded polypeptides in diagnosing colorectal
 PT cancer.
 XX Disclosure; Page 186; 266pp; English.
 PS
 XX The invention relates to a method of screening for nucleic acid molecules
 CC that show altered expression in a first cell sample comprising comparing
 CC the gene expression profile of the sample with that of a second reference
 CC sample, where the first sample has been grown in the presence of butyrate
 CC or a related carbon source from which butyrate is directly or indirectly
 CC derived, but the reference sample has not. The invention also relates to
 CC a method of detecting at least one nucleic acid molecule associated with
 CC the initiation and/or progression of colorectal cancer in an animal,
 CC comprising providing a biological sample comprising at least one cell to
 CC be tested, contacting the sample with a ligand (preferably a hybridizing
 CC nucleic acid molecule) which binds to at least one nucleic acid and
 CC detecting the presence of at least one molecule in the sample, a method
 CC of detecting at least one polypeptide associated with the initiation
 CC and/or progression of colorectal cancer in an animal comprising providing
 CC a biological sample comprising at least one cell to be tested, contacting
 CC the sample with at least one ligand that specifically binds at least one
 CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising
 CC an amino acid sequence which varies by the addition, deletion or
 CC substitution of at least one amino acid residue and detecting the
 CC presence of the polypeptide in the sample, a method of screening for
 CC agents that modulate the activity of at least one polypeptide encoded by
 CC a gene associated with the initiation and/or progression of colorectal
 CC cancer comprising forming a preparation comprising at least one
 CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising
 CC at least one addition, deletion or substitution and at least one agent to
 CC be tested and determining the activity of the agent with respect to
 CC activity of the polypeptide, and an antibody or its effective binding
 CC portion identified by the method, for use as a pharmaceutical. The
 CC methods are useful for screening for nucleic acid molecules that show
 CC altered expression in a cell sample, and for detecting a nucleic acid and
 CC a polypeptide respectively, that are associated with the initiation
 CC and/or progression of colorectal cancer and are useful for detecting or
 CC monitoring colorectal cancer, especially adenocarcinoma. The methods are
 CC also useful for screening for agents that modulate the activity of at
 CC least one polypeptide encoded by a gene associated with the initiation
 CC and/or progression of cancer, where agents identified by the method are
 CC useful for treating colorectal cancer. The methods could also be used to
 CC detect or monitor other conditions such as colitis, Crohn's disease or
 CC irritable bowel syndrome, as a screening tool for fiber consumption, as
 CC an assay for colon microflora functionality or for early detection of pre
 CC -cancerous growth. This sequence represents a human nucleic acid
 CC identified by the screening method of the invention.
 XX
 SQ Sequence 522 BP; 99 A; 171 C; 170 G; 82 T; 0 U; 0 Other;
 Query Match 94.7%; Score 74.8; DB 14; Length 522;
 Best Local Similarity 97.4%; Pred. No. 5.7e-14;
 Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 Db |||||||
 82 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 23
 |||||||
 QY 61 AACAAACCGACTCTAATA 78
 Db |||||||
 22 AACAAACCGACTCACA 5

RESULT 5

ADP81837
 ID ADF81837 standard; DNA; 592 BP.

XX AC ADF81837;
 XX DT 26-FEB-2004 (first entry)
 XX DE Leukaemia-related DNA sequence #2393.
 XX KW Cytostatic; Gene therapy; leukaemia; ss.
 XX OS Unidentified.

XX PN WO2003039443-A2.
 XX PD 15-MAY-2003.
 XX PF 04-NOV-2002; 2002WO-EP012303.
 XX PR 05-NOV-2001; 2001EP-00126244.
 XX PR 30-APR-2002; 2002EP-00009758.
 XX PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX PA (YLU-) UNIV LUDWIG MAXIMILIANS.
 XX PA (HAFE/) HAFERLACH T.
 XX PA (SCHO/) SCHOCH C.
 XX PA (KERN/) KERN W.

XX PI Haferlach T, Schoch C, Kern W, Kohlmann A, Schnittgger S, Dugas M;
 XX PI Ellis R, Brors B, Mergenthaler S;
 XX PS WPI; 2003-505037/47.
 XX PT Determining the subtype of leukemia cells and whether a patient sample
 XX PT contains leukemia cells or other cells, useful for treating leukemia,
 XX PT comprises determining the expression profile of a group of markers in a
 XX PT patient sample.
 XX PS Disclosure; SEQ ID NO 2393; 2938pp; English.
 XX CC The present invention relates to a method (M1) for determining the
 XX CC subtype of leukemia cells and whether a patient sample contains
 XX CC leukemia cells. The method comprises determining the expression profile
 XX CC of a group of markers in a patient sample. The method is useful for
 XX CC determining the presence of leukemia cells, its types or subtypes, and
 XX CC for the preparation of a medicament for treating leukemia.
 XX SQ Sequence 592 BP; 133 A; 173 C; 178 G; 108 T; 0 U; 0 Other;

Query Match 94.7%; Score 74.8; DB 10; Length 592;
 Best Local Similarity 97.4%; Pred. No. 5.9e-14;
 Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 Db |||||||
 434 CGGGAATCCCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 493
 |||||||
 QY 61 AACAAACCGACTCTAATA 78
 Db |||||||
 494 AACAAACCGACTCACA 511

RESULT 6

ADQ86642
 ID ADQ86642 standard; cDNA; 1985 BP.

XX AC ADQ86642;
 XX DT 07-OCT-2004 (first entry)
 XX DE Human tumour-associated antigenic target (TAT) cDNA sequence #3516.

XX KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
 XX KW cancer; cell proliferative disorder; gene; ss.
 XX OS Homo sapiens.
 XX PN WO2004060270-A2.
 XX PD 22-JUL-2004.
 XX PF 15-OCT-2003; 2003WO-US029126.
 XX PR 18-OCT-2002; 2002US-0418988P.

XX PA (GETH) GENENTECH INC.
 XX PA (WUTD/) WU T D.
 XX PA (ZHOU/) ZHOU Y.

XX PI Wu TD, Zhou Y;
 XX PS WPI; 2004-534300/51.

XX DR New nucleic acid molecule and encoded polypeptide, for diagnosing,
 XX PT preventing or treating cell proliferative disorders such as cancer.
 XX PS Claim 1; SEQ ID NO 3516; 5504pp; English.
 XX CC The present invention describes an isolated tumour-associated antigenic
 XX CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
 XX CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
 XX CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
 XX CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
 XX CC (c). Also described: (1) an expression vector comprising the above
 XX CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
 XX CC a process for producing a polypeptide; (4) an isolated polypeptide
 XX CC comprising: (a) an amino acid sequence encoded by any of the above
 XX CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
 XX CC length coding region of the above nucleotide sequences; or (c) a sequence
 XX CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
 XX CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
 XX CC an isolated antibody that binds to the above polypeptide; (7) a process
 XX CC for producing the antibody; (8) an isolated oligopeptide that binds to
 XX CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
 XX CC binding organic molecule that binds to the above polypeptide; (10) a
 XX CC composition of matter comprising the above (chimeric) polypeptide,
 XX CC antibody, oligopeptide or TAT binding organic molecule, in combination
 XX CC with a carrier; (11) an article of manufacture comprising a container and
 XX CC the composition of matter contained within the container; (12) methods of
 XX CC inhibiting the growth of a cell that expresses the above protein, where
 XX CC the growth of the cell is at least in part dependent upon a growth
 XX CC potentiating effect of the above protein; (13) a method of
 XX CC therapeutically treating a mammal having a cancerous tumour comprising
 XX CC cells that express the above protein; (14) a method of determining the
 XX CC presence of a protein in a sample suspected of containing the protein
 XX CC described above; (15) methods of diagnosing the presence of a tumour in a
 XX CC mammal; (16) a method for treating or preventing a cell proliferative
 XX CC disorder associated with increased expression or activity of the above
 XX CC protein; and (17) a method of binding an antibody, oligopeptide or
 XX CC organic molecule to a cell that expresses the protein described above.
 XX CC The TAT sequences have cytostatic activities, and can be used in gene
 XX CC therapy. The composition and methods are useful for diagnosing,
 XX CC preventing or treating cancer. The composition is also used for preparing
 XX CC a medicament for the therapeutic treatment or diagnostic detection of a
 XX CC cell proliferative disorder or cancer. The present sequence represents a
 XX CC human TAT cDNA sequence from the present invention.

XX SQ Sequence 1985 BP; 372 A; 649 C; 594 G; 370 T; 0 U; 0 Other;
Query Match 94.7%; Score 74.8; DB 12; Length 1985;
Best Local Similarity 97.4%; Pred. No. 7.4e-14;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1636 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1695
Qy 61 AACAAACCGACTCTAAAA 78
Db 1696 AACAAACCGACTCTACAAA 1713
RESULT 7
ADQ85525
ID ADQ85525 standard; cDNA; 1985 BP.
AC ADQ85525;
DT 07-OCT-2004 (first entry)
XX Human tumour-associated antigenic target (TAT) cDNA sequence #2339.
DE human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.
XX Homo sapiens.
OS WO2004060270-A2.
PN 22-JUL-2004.
XX 15-OCT-2003; 2003WO-US029126.
XX 18-OCT-2002; 2002US-0418988P.
XX (GETH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX Wu TD, Zhou Y;
XX WPI; 2004-534300/51.
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
XX Claim 1; SEQ ID NO 2339; 5504pp; English.
XX The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridizes to (a)-
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of

CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein in a
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX
SQ Sequence 1985 BP; 372 A; 649 C; 594 G; 370 T; 0 U; 0 Other;
Query Match 94.7%; Score 74.8; DB 13; Length 1985;
Best Local Similarity 97.4%; Pred. No. 7.4e-14;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1636 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1695
Qy 61 AACAAACCGACTCTAAAA 78
Db 1696 AACAAACCGACTCTACAAA 1713
RESULT 8
AAC88152
ID AAC88152 standard; cDNA; 2000 BP.
AC AAC88152;
DT 13-MAR-2001 (first entry)
XX Human K12 nucleotide sequence SEQ ID NO:3.
XX Human; CD7; K12; cognate ligand; cluster of differentiation; cancer;
KW identification; inhibiting T cell proliferation; HIV; infection;
KW activating natural killer cell proliferation; leukaemia; lymphoma;
KW sepsis; graft versus host disease; autoimmune disease; arthritis;
KW multiple sclerosis; rheumatoid arthritis; psoriatic arthritis; lupus;
KW scleroderma; psoriasis; atopic dermatitis; type I diabetes mellitus;
KW Hashimoto's thyroiditis; pernicious anaemia; Addison's disease; uveitis;
KW myasthenia gravis; psoriasis; Guillain-Barre Syndrome; Grave's disease;
KW systemic lupus erythematosus; dermatomyositis; asthma; eczema;
KW atopic dermatitis; contact dermatitis; eczematous dermatitis;
KW seborrhoeic dermatitis; rhinitis; ss.
XX Homo sapiens.
OS WO2000073333-A2.
XX 07-DEC-2000.
XX 26-MAY-2000; 2000WO-US014612.
XX 28-MAY-1999; 99US-0136450P.
XX (IMMV) IMMUNEX CORP.
XX Lyman SD, Fanslow WC;
XX WPI; 2001-061511/07.
DR P-PSDB; AAB36658.
XX Stimulating intracellular signaling of CD7 comprises contacting a cell

PT expressing CD7 with recombinant K12 protein, the cognate ligand of CD7,
 PT to inhibit T cell proliferation and/or activate natural killer cell
 PT proliferation.

XX Disclosure; Page 36-38; 42pp; English.

XX The present invention describes a method for stimulating (S) the
 CC intracellular signalling of CD (cluster of differentiation) 7 comprising
 CC contacting a cell that expresses CD7 with a recombinant K12 protein (I),
 CC the cognate ligand of CD7. (S) is useful for inhibiting T cell
 CC proliferation and/or activating NK (natural killer) cell proliferation
 CC and/or inducing NK toxicity in a mammal which involves administration of
 CC K12 protein. It is also used for treating HIV-1 infection, cancer (T cell
 CC leukaemia, acute lymphoid leukaemia, cutaneous T cell lymphoma),
 CC bacterial and viral infections, mediated by CD7. In the case of treating
 CC T cell leukaemia the soluble K12 protein is covalently attached to a
 CC toxin. A disease mediated by CD7 such as sepsis, graft versus host
 CC disease due to transplantation, autoimmune diseases, multiple sclerosis,
 CC arthritis, rheumatoid arthritis, psoriatic arthritis, scleroderma, lupus,
 CC psoriasis, atopic dermatitis, type I diabetes mellitus, Hashimoto's
 CC thyroiditis, pernicious anaemia, Addison's disease, myasthenia gravis,
 CC uveitis, psoriasis, Guillain-Barre Syndrome, Grave's disease, systemic
 CC lupus erythematosus and dermatomyositis, asthma, eczema, atypical
 CC dermatitis, contact dermatitis, other eczematous dermatitides, seborrheic
 CC dermatitis, and rhinitis is also treated by administering a K12
 CC antagonist (neutralising antibody). The present sequence encodes the
 CC human K12 protein, which is given in the exemplification of the present
 CC invention

XX Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match 94.7%; Score 74.8; DB 4; Length 2000;
 Best Local Similarity 97.4%; Pred. No. 7.5e-14;
 Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGAATCCCTTCAGTCTTTTAAAAAGTTCCATGACTCGAATATCTGAATGAAGAA 60
 DB 1645 CGGGAATCCCTTCAGTCTTTTAAAAAGTTCCATGACTCGAATATCTGAATGAAGAA 1704
 QY 61 AACAAACCGACTCAAAA 78
 DB 1705 AACAAACCGACTCAAAA 1722

RESULT 9

ID ABL65927 standard; DNA; 2000 BP.

XX ABL65927;

DT 15-MAY-2002 (first entry)

XX Lung cancer related gene sequence SEQ ID NO:4264.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytoskeletal; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; db.

XX Homo sapiens.

XX W0200194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US010838.

XX 05-JUN-2000; 2000US-0209473P.

XX 05-JUN-2000; 2000US-0209531P.

XX 18-SEP-2000; 2000US-0233133P.

XX 18-SEP-2000; 2000US-0233617P.

XX 20-SEP-2000; 2000US-0234009P.

PR 20-SEP-2000; 2000US-0234052P.
 PR 22-SEP-2000; 2000US-0234509P.
 PR 22-SEP-2000; 2000US-0234567P.
 PR 25-SEP-2000; 2000US-0234923P.
 PR 25-SEP-2000; 2000US-0234924P.
 PR 25-SEP-2000; 2000US-0235077P.
 PR 25-SEP-2000; 2000US-0235082P.
 PR 25-SEP-2000; 2000US-0235134P.
 PR 25-SEP-2000; 2000US-0235280P.
 PR 26-SEP-2000; 2000US-0235637P.
 PR 26-SEP-2000; 2000US-0235638P.
 PR 27-SEP-2000; 2000US-0235711P.
 PR 27-SEP-2000; 2000US-0235720P.
 PR 27-SEP-2000; 2000US-0235840P.
 PR 27-SEP-2000; 2000US-0235863P.
 PR 28-SEP-2000; 2000US-0236028P.
 PR 28-SEP-2000; 2000US-0236032P.
 PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.
 PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237295P.
 PR 02-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 01-NOV-2000; 2000US-0237608P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 PA (AVAL-) AVALON PHARM.

Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 Soppet DR, Weaver Z;

WPI; 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a chemical
 agent to be tested for anti-neoplastic activity, and determining a change
 in expression of a gene of a signature gene set.

Claim 1; SEQ ID NO 4264; 44pp; English.

The present invention describes a method (M1) for screening for an anti-
 neoplastic agent. The method involves exposing cells to a chemical agent
 to be tested for anti-neoplastic activity, determining a change in
 expression of at least one gene (I) of a signature gene set, where (I)
 comprises a sequence (S) selected from 847 sequences given in ABL61664
 to ABL70110, or is at least 95% identical to (S), where a change in
 expression is indicative of anti-neoplastic activity. (I) has cytostatic
 activity and can be used in gene therapy. M1 can be used for screening an
 anti-neoplastic agent, and can be used for producing a product which is
 the data collected with respect to the anti-neoplastic agent as a result
 of M1, and the data is sufficient to convey the chemical structure and/or
 properties of the agent. M1 can be used in the treatment of cancer such
 as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
 prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
 cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
 cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
 tumour

Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match 94.7%; Score 74.8; DB 6; Length 2000;
 Best Local Similarity 97.4%; Pred. No. 7.5e-14;
 Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 Db 1645 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCGACTCAAAA 78
 |||||
 Db 1705 AACAAACCGACTCAAAA 1722

RESULT 10 ABK84431

ID ABK84431 standard; cDNA; 2000 BP.

XX AC ABK84431;

XX DT 14-AUG-2002 (first entry)

XX DE Human cDNA differentially expressed in granulocytic cells #1002.

XX KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.

XX OS Homo sapiens.

XX FN WO200228999-A2.

XX PD 11-APR-2002.

XX PF 03-OCT-2001; 2001WO-US030821.

XX PR 03-OCT-2000; 2000US-0237189P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Beazer-Barclay Y, Weisman SM, Yamaga S, Vockley J;

XX DR WPI; 2002-435328/46.

XX PT Detecting granulocyte activation by detecting differential expression of
 PT genes associated with granulocyte activation, which serves as diagnostic
 PT markers that is useful for monitoring disease states and drug toxicity.

XX PS Claim 1; SEQ ID NO 1002; 114pp; English.

XX CC The invention relates to detecting (M1) granulocyte (GC) activation
 CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing the
 CC expression level to an expression level in an unactivated GC, where
 CC differential expression of Gs is indicative of GCA. Also included are
 CC modulating (M2) CA by contacting GC with an agent that alters the
 CC expression of at least one gene in Gs; (2) screening (M3) for an agent
 CC capable of modulating GCA or an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease using the gene expression
 CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
 CC tissue, an allergic response in a subject, exposure of a subject to a
 CC pathogen or sterile inflammatory disease, by detecting the level of
 CC expression in a sample of the tissue of gene(s) from Gs, where the level
 CC of expression of the gene is indicative of inflammation; (4) treating
 CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease, by contacting a tissue having inflammation with an
 CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
 CC is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful
 CC for screening an agent capable of modulating GCA preferably in an
 CC inflammation in a tissue; M4 is useful for detecting an inflammation

CC (especially chronic) in a tissue, an allergic response in a subject,
 CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
 CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
 CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
 CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
 CC disease, ulcerative colitis, periodontal disease; also bacterial
 CC infection, viral infection, parasitic infection, protozoal infection,
 CC fungal infection and M5 is useful for treating one of the above
 CC conditions. The present sequence represents a gene differentially
 CC expressed in granulocytes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match 94.7%; Score 74.8; DB 6; Length 2000;
 Best Local Similarity 97.4%; Pred. No. 7.5e-14;
 Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 Db 1645 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCGACTCAAAA 78
 |||||

Db 1705 AACAAACCGACTCAAAA 1722

RESULT 11 ADJ74845

ID ADJ74845 standard; DNA; 2000 BP.

XX AC ADJ74845;

XX DT 20-MAY-2004 (first entry)

XX DE Marker gene SEQ ID NO:97.

XX KW bronchial asthma; chronic obstructive pulmonary disease;
 KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
 KW gene therapy; marker gene; gene; ds.

XX OS Homo sapiens.

XX PN EPI394274-A2.

XX PD 03-MAR-2004.

XX PF 04-AUG-2003; 2003EP-00254857.

XX PR 06-AUG-2002; 2002JP-00229312.

XX PR 20-MAR-2003; 2003JP-00077212.

XX PA (GENO-) GENOX RES INC.

XX PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuohara K;

XX DR WPI; 2004-193155/19.

XX CC Testing for bronchial asthma or chronic obstructive pulmonary disease by
 CC comparing the expression level of a marker gene in a biological sample
 CC from a subject with the expression level of the gene in a sample from a
 CC healthy subject.

XX PS Claim 1; SEQ ID NO 97; 241pp; English.

XX CC The present invention describes a method of testing for bronchial asthma
 CC or chronic obstructive pulmonary disease. The method comprises
 CC determining the expression level of a marker gene in a biological sample
 CC from a subject, comparing the expression level determined with the
 CC expression level of the marker gene in a biological sample from a healthy
 CC subject, and judging whether the subject has bronchial asthma or chronic

obstructive pulmonary disease. The marker gene comprises: (a) a group of genes (S1) whose expression levels increase when respiratory epithelial cells are stimulated with interleukin-13; or (b) a group of genes (S2) whose expression levels decrease when respiratory epithelial cells are stimulated with interleukin-13. Also described: (1) a reagent (I) for testing for bronchial asthma or chronic obstructive pulmonary disease; (2) a kit for screening for a candidate compound for a therapeutic agent to treat bronchial asthma or chronic obstructive pulmonary disease; (3) an animal model for bronchial asthma or chronic obstructive pulmonary disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a method for producing an animal model for bronchial asthma or chronic obstructive pulmonary disease; (6) a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease, comprising the compound, a marker gene or an antisense nucleic acid corresponding to a portion of the marker gene, a ribozyme, a polynucleotide that suppresses the expression of the gene through an RNAi effect or an antibody recognising a protein encoded by a marker gene; and (7) a DNA chip for testing for bronchial asthma or a chronic obstructive pulmonary disease, on which a probe has been immobilised to assay a marker gene. (I) has respiratory and antihistaminic activities, and can be used in gene therapy. The method is useful for testing for or screening for a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease. The present sequence is used in the exemplification of the present invention.

Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other

SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;
Query Match 94.7%; Score 74.8; DB 12; Length 2000;
Best Local Similarity 97.4%; Pred. No. 7.5e-14;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	CGGGAATCCCCCTTCAGTCTCTTTGAAAAGAGTTCCATGACTCGAATATCTCTGAAATGAAGAA	60
Db	1645	CGGGAATCCCCCTTCAGTCTCTTTGAAAAGAGTTCCATGACTCGAATATCTCTGAAATGAAGAA	1704
Qy	61	AACAAACCGGACTCTAAAA	78
Db	1705	AACAAACCGGACTCTACAA	1722

RESULT 12
ADR52746
ID ADR52746 standard; DNA; 2000 BP.

DT 18-NOV-2004 (first entry)

Drug therapy altered expressed gene #97.

drug activity monitoring; expression profile; gene expression;
peripheral blood sample; peripheral blood mononuclear cell; drug therapy;
CCI-779; immunosuppressant; rapamycin; mammalian target of rapamycin;
mTOR; ds.

XX
OS
Homo sapiens.

AA
PN
WO2004072265-A2.

26-AUG-2004.

11-FEB-2004: 2004WO-US004118.

XX
PR 11-FEB-2003; 2003US-0446133P.

PR 03-APR-2003; 2003US-0459782P;
PR 23-JAN-2004; 2004US-0538246P;

XX
PA (AMHP) WYETH.

PA (BURC/) BURCZY
PA (TWIN/) TWINE

PA (DORN//) DORNER A J.

XX

FT BUTCHER M, TWINE N,

XX
DR WPI; 2004-642301/62.

Monitoring drug activities in vivo comprises comparing an expression profile of a gene in a peripheral blood sample of a patient before and after drug therapy.

PS Disclosure; SEO ID NO 97; 136pp; English.

The invention relates to a method of monitoring drug activities in vivo by comparing an expression profile of at least one gene in a peripheral blood sample of a patient to a reference expression profile of the at least one gene, where the at least one gene is differentially expressed in peripheral blood mononuclear cells (PBMCs) of patients who have a non-blood disease and are subjected to a drug therapy as compared to PBMCs isolated from the patient before the drug therapy, and where the patient has the non-blood disease and is being treated by the drug therapy. The method, kit, and nucleic acid array are useful for monitoring drug activities in vivo. The drug is especially CGI-779, an ester analogue of the immunosuppressant rapamycin which is a potent inhibitor of the mammalian target of rapamycin (mTOR). This sequence represents a gene expressed in PBMC altered by the drug therapy. (Note: this sequence does not form part of the printed specification but was obtained in electronic format from WIPO at ftp.wipo.int/pub/published/pct/sequences/).

Sequence 2000 BP: 367 A: 659 C: 602 G: 372 T: 0 U: 0 Other: 0

Query Match	Score	DB 13;	Length
Best Local Similarity	94.7%	74.8;	2000;
Matches	97.4%;		
76: Conservative	Pred. No. 7.5e-14;		
0: Mismatches	0;		
Indels	2;		
Gaps	0;		

Qy	1	CGGGAATCCCCCTTCAGTCTTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA	60
Dp	1645	CGGGAATCCCCCTTCAGTCTTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA	1704

61 AACAAACCGACTCTAAA 78
Ov

RESULT 13

ADP24663
ID ADP24663 standard; CDNA; 2000 BP.

XX
AC ADP24663;

18-NOV-2004 (first entry)

XX
DE
PRO polypeptide encoding cDNA SEQ ID NO:1841.

ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;
 KW immunosuppressive; osteopathic; antidiabetic; dermatological;
 KW antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;
 KW gene therapy; immune system.

XX OS Unidentified.

XX
PN
WO2004041170-A2.

XX PD 21-MAY-2004.

XX
PF 30-OCT-2003: 2003WO-US034312.XX
PR 01-NOV-2002: 2002US-0423394P.

XX PA (GETH) GENENTECH INC.

XX Clark H, Schoenfeld J, van Lookeren M, Williams PM, Wood WI;
PI PI Wu TD;
PI

XX WPI; 2004-419628/39.
DR P-PSDB; ADP24664.
DR

XX

PT New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.

XX Claim 1; SEQ ID NO 1841; 2940pp; English.

XX The invention relates to a novel isolated nucleic acid and the PRO
XX polypeptide encoded by it. A protein of the invention has
XX antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
XX osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
XX antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
XX of the invention may have a use in gene therapy. The PRO polypeptide, its
XX agonist, antagonist, or antibody that specifically binds to the
XX polypeptide is useful for treating an immune related disorder such as
XX systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
XX juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
XX idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
XX vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
XX thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
XX disease, a demyelinating disease of the central or peripheral nervous
XX system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
XX a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
XX disease, infectious or autoimmune chronic active hepatitis, primary
XX biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
XX inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
XX disease, an autoimmune or immune-mediated skin disease, a bullous skin
XX disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
XX disease, asthma, allergic rhinitis, atopic dermatitis, food
XX hypersensitivity, urticaria, an immunologic disease of the lung,
XX eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
XX pneumonitis, a transplantation associated disease, graft rejection or
XX graft-versus-host disease. The present sequence encodes a PRO protein of
XX the invention.

XX Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match 94.7%; Score 74.8; DB 13; Length 2000;
Best Local Similarity 97.4%; Pred. No. 7.5e-14;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGGAATCCCTTCAGTCTTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGGAATCCCTTCAGTCTTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704
Qy 61 AACAAACCGACTCTCAAAA 78
Db 1705 AACAAACCGACTCTCAAAA 1722

RESULT 14
ADU06261
ID ADU06261 standard; DNA; 2000 BP.

XX ADU06261;

XX 27-JAN-2005 (first entry)

XX Novel bronchial cancer-associated human gene SeqID485.

XX bronchial cancer; cytostatic; tumour-associated protein;
XX cancer detection; metastasis; tumour; gene; ds; human.

XX Homo sapiens.

XX DE10316701-A1.

XX 04-NOV-2004.

XX 09-APR-2003; 2003DE-01016701.

XX 09-APR-2003; 2003DE-01016701.

PA (HINZ/) HINZMANN B.
PA (HERM/) HERMANN K.
XX (CAST/) HEIDEN CASTANOS-VELEZ E.

PI Mennerich D, Bruemendorf T, Heiden E, Hermann K, Kinnemann H;
PI Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pflarsky C;

XX WPI; 2004-786403/78.

XX P-PSDB; ADU06748.

XX New nucleic acid, and derived proteins, useful for diagnosis of bronchial
XX cancer and in screening for therapeutic and diagnostic agents.

XX Claim 1; SEQ ID NO 485; 1381pp; German.

XX This invention relates to a novel isolated nucleic acid associated with
XX bronchial cancer comprising 489 defined sequences given in the
XX specification. The invention may be useful for the production of
XX compounds with a cytostatic activity through the inhibition of expression
XX or activity of tumour-associated proteins. The novel DNA sequences and
XX the proteins/peptides encoded by them are used for detecting bronchial
XX cancer or determining the risk of developing it and to screen for
XX specific binding partners of the DNA or protein sequences, where the
XX binding partners are potentially useful as agents for treating or
XX diagnosing bronchial cancer. The DNA or protein sequences can also be
XX used for prognosis, detection of metastases and for secondary treatment
XX (of tumours that have been stabilised or are no longer detectable).
XX Detecting abnormal expression of the DNA sequences provides early
XX diagnosis of bronchial cancers. The present sequence is that of a novel
XX bronchial cancer-associated human gene sequence of the invention.

XX Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match 94.7%; Score 74.8; DB 13; Length 2000;
Best Local Similarity 97.4%; Pred. No. 7.5e-14;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGGAATCCCTTCAGTCTTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704
Qy 61 AACAAACCGACTCTCAAAA 78
Db 1705 AACAAACCGACTCTCAAAA 1722

RESULT 15
ADW38555
ID ADW38555 standard; cDNA; 2000 BP.

XX ADW38555;

XX 24-MAR-2005 (first entry)

XX Immunomodulatory gene SECTM1 cDNA.

XX cytostatic; immunosuppressive; virucide; diagnosis; prognosis;
XX pharmaceutical; immunotherapy; cancer; cytostatic; neoplasm;
XX autoimmune disease; immunosuppressive; immune disorder; viral infection;
XX infection; SECTM1; gene; ss.

XX Homo sapiens.

XX WO2005000099-A2.

XX 06-JAN-2005.

XX 09-JUN-2004; 2004WO-US018461.

XX 09-JUN-2003; 2003US-0477291P.

XX (GENZ) GENZYME CORP.

PI Roberts BL;
XX
DR MPI; 2005-058046/06.
DR P-PSDB; ADM38554.
XX
PT New isolated blood factor domain polynucleotides having immunomodulatory
PT activity, useful for aiding in the diagnosis or treating disorders
PT relating to the immune responses, e.g. cancers, autoimmune diseases, or
PT viral infections.
XX
PS Disclosure; SEQ ID NO 64; 141pp; English.
XX
CC The invention describes an isolated polynucleotide (I) encoding a peptide
CC selected from 21 polynucleotides fully given in the specification, or the
CC complement of the polynucleotide. Also described are: an isolated peptide
CC selected from 21 peptides fully given in the specification; a host cell
CC comprising (i) or the peptide; an antibody that specifically recognizes
CC and binds the peptide; a composition comprising the host cell, the
CC peptide, or the antibody, and a pharmaceutical carrier; an immune
CC effector cell raised in the presence and at the expense of a host cell;
CC and a method for eliciting a cytolytic response in a subject. Also
CC disclosed are: a method for monitoring gene expression; a method for
CC modulating the expression of the immunomodulatory polynucleotides and
CC expression products; a method for screening for candidate agents that
CC modulate the expression of the polynucleotide or the expression products
CC of the polynucleotide; assays for the identification, assessment, and
CC development of candidate agents capable of modulating the activity of the
CC polynucleotides or polypeptides; a method for monitoring an immune
CC response in a subject; and a method for active immunotherapy. The
CC polynucleotides (e.g., blood factor domains) having immunomodulatory
CC activity are useful for detecting, diagnosing, prognosing, or monitoring
CC the progression of a disease. They are useful for aiding in the diagnosis
CC or treating disorders relating to the immune responses, e.g. cancers,
CC autoimmune diseases, or viral infections. This sequence encodes
CC immunomodulatory gene SECTM1.
XX
SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match 94.7%; Score 74.8; DB 14; Length 2000;
Best Local Similarity 97.4%; Pred. No. 7.5e-14;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCGCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAATGAGAA 60
Db 1645 CGGGAATCCCGCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAATGAGAA 1704

Qy 61 AACAAACCGACTCTAAAA 78
Db 1705 AACAAACCGACTCAGAA 1722

Search completed: December 7, 2005, 00:03:00
Job time : 127.194 secs

QY 61 AACAAACCGACTCTATAAA 78
|||||
Db 1705 AACAAACCGACTCACAAA 1722

RESULT 2

US-09-997-165-3
; Sequence 3, Application US/09997165
; Patent No. US20020141999A1
; GENERAL INFORMATION:
; APPLICANT: Lyman, Stewart D.
; APPLICANT: Fanelow, William C.
; TITLE OF INVENTION: LIGAND FOR CD7 AND METHODS OF USE THEREOF
; FILE REFERENCE: 2913-US
; CURRENT APPLICATION NUMBER: US/09/997,165
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: PCT/US00/14612
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/136,450
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)..(865)
US-09-997-165-3

Query Match 94.7%; Score 74.8; DB 3; Length 2000;
Best Local Similarity 97.4%; Pred. No. 1.1e-13;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCATGACTCGAATATCTCGAATGAAGAA 60
|||||
Db 1645 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCATGACTCGAATATCTCGAATGAAGAA 1704
|||||

QY 61 AACAAACCGACTCTATAAA 78
|||||
Db 1705 AACAAACCGACTCACAAA 1722

RESULT 3

US-10-775-169-97
; Sequence 97, Application US/10775169
; Publication No. US2004017543A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dörner, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 97
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-169-97

Query Match 94.7%; Score 74.8; DB 8; Length 2000;
Best Local Similarity 97.4%; Pred. No. 1.1e-13;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCATGACTCGAATATCTCGAATGAAGAA 60
|||||

Db 1645 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCATGACTCGAATATCTCGAATGAAGAA 1704
QY 61 AACAAACCGACTCTATAAA 78
|||||
Db 1705 AACAAACCGACTCACAAA 1722

RESULT 4

US-10-843-641A-4264
; Sequence 4264, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: Signature Gene Sets
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4264
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-4264

Query Match 94.7%; Score 74.8; DB 9; Length 2000;
Best Local Similarity 97.4%; Pred. No. 1.1e-13;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCATGACTCGAATATCTCGAATGAAGAA 60
|||||
Db 1645 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCATGACTCGAATATCTCGAATGAAGAA 1704
|||||

QY 61 AACAAACCGACTCTATAAA 78
|||||
Db 1705 AACAAACCGACTCACAAA 1722

RESULT 5

US-10-631-467-97
; Sequence 97, Application US/10631467
; Publication No. US20050208496A1
; GENERAL INFORMATION:
; APPLICANT: Genox Research Inc.
; TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive p
; FILE REFERENCE: 3462.1005-000
; CURRENT APPLICATION NUMBER: US/10/631,467
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: JP 2003-077212
; PRIOR FILING DATE: 2003-03-20

;; PRIOR APPLICATION NUMBER: JP 2002-229312
;; PRIOR FILING DATE: 2002-08-06
;; NUMBER OF SEQ ID NOS: 2086
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 97
;; LENGTH: 2000
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-631-467-97

Query Match 94.7%; Score 74.8; DB 9; Length 2000;
Best Local Similarity 97.4%; Pred. No. 1.1e-13;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704
Qy 61 AACAAACCGACTCTAAAA 78
Db 1705 AACAAACCGACTCACAA 1722

RESULT 6

US-10-080-522-2
;; Sequence 2, Application US/10080522
;; Publication No. US20030096326A1
;; GENERAL INFORMATION:
;; APPLICANT: KAUFMAN, RUSSEL E.
;; SLENTZ-KESLER, KIMBERLY
;; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
;; CELLS
;; NUMBER OF SEQUENCES: 2
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: NIXON & VANDERHYE P.C.
;; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
;; CITY: ARLINGTON
;; STATE: VIRGINIA
;; COUNTRY: U.S.A.
;; ZIP: 22201-4714

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/10/080,522
;; FILING DATE: 25-Feb-2002
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 09/539,774
;; FILING DATE: 31-MAR-2000
;; APPLICATION NUMBER: US 09/210,474
;; FILING DATE: 14-DEC-1998
;; APPLICATION NUMBER: US 08/755,559
;; FILING DATE: 22-NOV-1996

;; ATTORNEY/AGENT INFORMATION:
;; NAME: WILSON, MARY J.
;; REGISTRATION NUMBER: 32,955
;; REFERENCE/DOCKET NUMBER: 1579-645
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 816-4000
;; TELEFAX: (703) 816-4100
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2180 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-080-522-2

Query Match 94.7%; Score 74.8; DB 5; Length 2180;
Best Local Similarity 97.4%; Pred. No. 1.1e-13;
Matches 76; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1841 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1900
Qy 61 AACAAACCGACTCTAAAA 78
Db 1901 AACAAACCGACTCACAA 1918

RESULT 7

US-10-437-963-1665/c
;; Sequence 1665, Application US/10437963
;; Publication No. US2004012343A1
;; GENERAL INFORMATION:
;; APPLICANT: La Rosa, Thomas J.
;; APPLICANT: Kovalic, David K.
;; APPLICANT: Zhou, Yihua
;; APPLICANT: Cao, Yongwei
;; APPLICANT: Wu, Wei
;; APPLICANT: Boukharov, Andrey A.
;; APPLICANT: Barbazuk, Brad
;; APPLICANT: Li, Ping
;; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;; FILE REFERENCE: 38-21(53221)B
;; CURRENT APPLICATION NUMBER: US/10/437,963
;; CURRENT FILING DATE: 2003-05-14
;; NUMBER OF SEQ ID NOS: 204966
;; SEQ ID NO 1665
;; LENGTH: 1166
;; TYPE: DNA
;; ORGANISM: Oryza sativa
;; FEATURE:
;; OTHER INFORMATION: Clone ID: PAT_MRT4530_101509C.1
US-10-437-963-1665

Query Match 38.5%; Score 30.4; DB 7; Length 1166;
Best Local Similarity 67.2%; Pred. No. 16;
Matches 43; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 13 TTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAACCAACCGACT 72
Db 1149 TTTTCTTTTATTAAGCAAGCTCAGTCATGATCTGAATTGAAGCAACCAACCGAAT 1090
Qy 73 CTAA 76
Db 1089 CTGA 1086

RESULT 8

US-10-398-221-3432
;; Sequence 3432, Application US/10398221
;; Publication No. US20040018514A1
;; GENERAL INFORMATION:
;; APPLICANT: KUNST, Frederik
;; APPLICANT: GLASER, Philippe
;; TITLE OF INVENTION: Listeria innocua, genome and applications
;; FILE REFERENCE: 344 702 - US
;; CURRENT APPLICATION NUMBER: US/10/398,221
;; CURRENT FILING DATE: 2003-03-27
;; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
;; PRIOR FILING DATE: 2001-10-04
;; PRIOR APPLICATION NUMBER: FR 00/12 697
;; PRIOR FILING DATE: 2000-10-04
;; NUMBER OF SEQ ID NOS: 4025
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 3432
;; LENGTH: 2278
;; TYPE: DNA

```
; ORGANISM: Listeria monocytogenes 4b
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u
; US-10-398-221-3432

Query Match          37.0%; Score 29.2; DB 7; Length 2278;
Best Local Similarity 62.2%; Pred. No. 48;
Matches 46; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 5 AATCCCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAAACA 64
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1274 AATCCCCCTTCATGCTAAACAAACACTTTATTGATTAGATAAAGCATACCATATAAAGA 1333

Qy 65 AACCGACTCTAAAA 78
    ||| |||||
Db 1334 AAGCAGGGATAAAA 1347

RESULT 9
US-09-814-353-17505/c
; Sequence 17505, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17505
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-17505

Query Match          36.7%; Score 29; DB 3; Length 528;
Best Local Similarity 61.0%; Pred. No. 36;
Matches 47; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 2 GGGAAATCCCCCTTCAGTCTTTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAAA 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GGGAAATTTCTATGCAGGGCTTTAAAAAAAATAAGTGATTTCGATTTTAAAAAAAATAA 122

Qy 62 ACAACCGACTCTAAAA 78
    ||||| |||||
Db 121 AAAAAAAAATAAATAA 105

RESULT 10
US-09-925-065A-787106/c
; Sequence 787106, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```

```
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 787106
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-787106

Query Match          36.5%; Score 28.8; DB 4; Length 582;
Best Local Similarity 69.6%; Pred. No. 43;
Matches 39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 6 ATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAA 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 179 ATGCCTCTCTGTATTATTTAAAAAAGTTATTTCATTAGAGTATCAGATTGAAAAGA 124

RESULT 11
US-09-925-065A-787107/c
; Sequence 787107, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 787107
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-787107

Query Match          36.5%; Score 28.8; DB 4; Length 582;
Best Local Similarity 69.6%; Pred. No. 43;
Matches 39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 6 ATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAA 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 179 ATGCCTCTCTGTATTATTTAAAAAAGTTATTTCATTAGAGTATCAGATTGAAAAGA 124

RESULT 12
US-09-925-065A-695404/c
; Sequence 695404, Application US/09925065A
```

```
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 695404
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-925-065A-695404

Query Match      36.5%; Score 28.8; DB 4; Length 1901;
Best Local Similarity 65.6%; Pred. No. 62;
Matches 42; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 13 TTCAGTCTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAAAAACCAACCGACT 72
   |||||
Db 343 TTCATTCTTTGAAATTTTCTTGACTCTATGAATAAAGTGATTAAGCCACAGAAA 284

QY 73 CTAA 76
   |||
Db 283 CAAA 280
```

```
RESULT 13
US-10-087-192-736
; Sequence 736, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 736
; LENGTH: 31828
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-087-192-736
```

```
Query Match      36.5%; Score 28.8; DB 5; Length 31828;
Best Local Similarity 69.6%; Pred. No. 1.4e+02;
Matches 39; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 6 ATCCCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAAA 61
   |||||
Db 20828 ATGCCTCTCTGTATTTAAAAAAGTTATTTTCATTAGAGATATCAGAAATGAAGAAA 20883
```

RESULT 14

```
US-10-312-841-1/c
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1
```

```
Query Match      36.2%; Score 28.6; DB 6; Length 3673778;
Best Local Similarity 64.2%; Pred. No. 5.9e+02;
Matches 43; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 13 TTCAGTCTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAAAAACCAACCGACT 72
   |||||
Db 3131340 TCCATATTTTTCAAACTAATACTATATTAATAAATACGAAAAAAAACCTCCT 3131281

QY 73 CTAAAAA 79
   |||||
Db 3131280 TTAATAA 3131274
```

```
RESULT 15
US-09-925-065A-695403/c
; Sequence 695403, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 695403
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-925-065A-695403
```

```
Query Match      35.9%; Score 28.4; DB 4; Length 1901;
Best Local Similarity 64.1%; Pred. No. 83;
Matches 41; Conservative 1; Mismatches 22; Indels 0; Gaps 0;

QY 13 TTCAGTCTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAAAAACCAACCGACT 72
   |||||
Db 343 TTCATTCTTTGAAATTTTCTTGACTCTATGAATAAAGTGATTAAGCCACAGAAA 284

QY 73 CTAA 76
   |||
```

Db 283 CAAA 280

Search completed: December 6, 2005, 22:56:59
Job time : 201.712 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:34:11 ; Search time 53.9271 Seconds
(without alignments)
455.919 Million cell updates/sec

Title: US-09-980-046B-3

Perfect score: 79

Sequence: 1 cgggaatcccccttcagttc.....aaacaaacgactctaaaaa 79

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA New:*

- 1: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	27.4	34.7	1131	US-10-750-185-51109 Sequence 51109, A
2	26.6	33.7	2995	US-10-750-185-35627 Sequence 35627, A
3	25.8	33.2	999	US-10-750-185-41906 Sequence 41906, A
4	25.8	32.7	2933	US-10-750-185-38312 Sequence 38312, A
5	25.4	32.2	3814	US-10-750-185-28705 Sequence 28705, A
6	25.2	31.9	490	US-11-108-172-317 Sequence 317, App
7	25.2	31.9	1489	US-10-750-185-28550 Sequence 28550, A
8	25	31.6	837	US-10-750-185-46504 Sequence 46504, A
9	25	31.6	1872	US-10-750-185-52067 Sequence 52067, A
10	25	31.6	2066	US-10-523-588-2 Sequence 2, Appli
11	24.8	31.4	600	US-10-750-185-3989 Sequence 3989, Ap
12	24.8	31.4	639	US-10-793-626-1627 Sequence 1627, Ap
13	24.8	31.4	639	US-10-793-626-1965 Sequence 1965, Ap
14	24.8	31.4	1274	US-10-750-185-29440 Sequence 29440, A
15	24.8	31.4	1693	US-10-750-185-34327 Sequence 34327, A
16	24.8	31.4	3227	US-10-793-626-3688 Sequence 3688, Ap
17	24.8	31.4	3227	US-10-793-626-3387 Sequence 3387, Ap
18	24.8	31.4	4373	US-10-750-185-33226 Sequence 33226, A
19	24.8	31.4	186882	US-11-112-908-23 Sequence 23, Appli
20	24.6	31.1	1026	US-10-750-185-48087 Sequence 48087, A
21	24.6	31.1	1662	US-10-750-185-32969 Sequence 32969, A
22	24.4	30.9	727	US-10-750-185-60268 Sequence 60268, A
23	24.4	30.9	1079	US-10-750-185-29189 Sequence 29189, A

ALIGNMENTS

RESULT 1

US-10-750-185-51109
; Sequence 51109, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51109
; LENGTH: 1131
; TYPE: DNA
; ORGANISM: Bovine 19866880844238
US-10-750-185-51109

Query Match 34.7%; Score 27.4; DB 6; Length 1131;

Best Local Similarity 65.6%; Pred. No. 3.7;

Matches 40; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 18 TTCTTTGAAAAGTTTCATGACTGATATCTGAATGAGAAAACCGACTCTAAA 77

Db 995 TTGCTAAATAATTTTCAATGATGCTTTATTTCTAAAAAAGAAAAACCAATTATGAG 1054

QY 78 A 78

Db 1055 A 1055

RESULT 2

US-10-750-185-35627
; Sequence 35627, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.


```
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52067
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Bovine 19866880381128
US-10-750-185-52067

Query Match      31.6%; Score 25; DB 6; Length 1872;
Best Local Similarity 61.5%; Pred. No. 25;
Matches 40; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 13 TTCAAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAAACAAACCGACT 72
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1464 TTAGTTCTTTTATATCTTCCTGCAATTTCTTAAATCTGGACAGAGAAATCTAGGCAAC 1523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 73 CTAAG 77
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1524 CCACA 1528

RESULT 10
US-10-523-588-2
; Sequence 2, Application US/10523588
; Publication No. US20050251870A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: CSNK1GS AS MODIFIERS OF THE P21 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX03-057C-US
; CURRENT APPLICATION NUMBER: US/10/523,588
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: US 60/401,739
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 2066
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-523-588-2

Query Match      31.6%; Score 25; DB 6; Length 2066;
Best Local Similarity 58.9%; Pred. No. 25;
Matches 43; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 7 TCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAAACAA 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1953 TGCACTTTATAGTTTCTTTTAAAAAACAACAAAAACATGGCAGATGCACAAAAA 2012
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 67 CCGACTCTAAAAA 79
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2013 AAAAAA 2025

RESULT 11
US-10-750-185-3989
; Sequence 3989, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
```

```
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3989
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT01900
US-10-750-185-3989

Query Match      31.4%; Score 24.8; DB 6; Length 600;
Best Local Similarity 57.9%; Pred. No. 21;
Matches 44; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 4 GAATCCCTTTCAGTTCTTTTGA AAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAC 63
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 310 GAGTCTCCATTCATTGCTCAATAACCAAGATCACACTACTTTCTTAATTCACATAAA 369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 64 AAACGACTCTAAAAA 79
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 370 GACCTATCTCTAAACA 385

RESULT 12
US-10-793-626-1627
; Sequence 1627, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1627
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-1627

Query Match      31.4%; Score 24.8; DB 6; Length 639;
Best Local Similarity 63.3%; Pred. No. 22;
Matches 38; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 20 CTTTGA AAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAAACAAACCGACTCTAAAAA 79
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152 CTGTGATAGATTTCACTGACTCAAGTTTTCAGTTCAGTCATCAAGAGGACTGAAAAACA 211

RESULT 13
US-10-793-626-1965
; Sequence 1965, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1965
; LENGTH: 639
```

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-1965

Query Match      31.4%; Score 24.8; DB 6; Length 639;
Best Local Similarity 63.3%; Pred. No. 22;
Matches 38; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 20 CTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAACAAACCGACTCTTAAAAA 79
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152 CTGATAGATATTTCATCGACTCAAGTTTTTTCAGTTCAAGTCATCAAGGGGACTGAAAAACA 211

RESULT 14
US-10-750-185-29440
; Sequence 29440, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 29440
; LENGTH: 1274
; TYPE: DNA
; ORGANISM: Bovine 19866880429041
US-10-750-185-29440

Query Match      31.4%; Score 24.8; DB 6; Length 1274;
Best Local Similarity 57.9%; Pred. No. 26;
Matches 44; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 4 GAATCCCCCTTCAGTCTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAAC 63
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 655 GAGTCTCCATTTCATTGCTCAAATACCAAGAAATCACACTACTTTTCTTAATTCAACATAAA 714

Qy 64 AAACCGACTCTAAAAA 79
Db 715 GACCTATCTCTAAAAACA 730

RESULT 15
US-10-750-185-34327
; Sequence 34327, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
```

```
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 34327
; LENGTH: 1693
; TYPE: DNA
; ORGANISM: Bovine 19866881065593
US-10-750-185-34327

Query Match      31.4%; Score 24.8; DB 6; Length 1693;
Best Local Similarity 63.3%; Pred. No. 28;
Matches 38; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 14 TCAGTCTTTTGAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAAACAAACCGACTC 73
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1565 TCAGTCTTTTGGAGAGTGCATCACCTTAAACCGAAGATGAAAAATTTCTAATAAGTCTC 1624

Search completed: December 6, 2005, 23:16:50
Job time : 53.9271 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 139.915 Seconds
(without alignments)
4239.411 Million cell updates/sec

Title: US-09-980-046B-4
Perfect score: 89
Sequence: 1 cgggaatcccccttcagtc.....aaaaaaaaaaaaaagaaaaa 89

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_21.*

1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2003ds.*
12: geneseqn2004as.*
13: geneseqn2004bs.*
14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	89	100.0	90	4	AAC89735
2	79.4	89.2	1464	3	AAC59772 Human sec
3	71	73.8	79	4	AAC89734 Human sec
4	70	78.7	428	13	ADU14748 Solid tum
5	70	78.7	522	14	AEA36210 Human nuc
6	70	78.7	592	10	ADF81837 Leukaemia
7	70	78.7	1985	12	ADQ86642 Human tum
8	70	78.7	1985	13	ADQ85525 Human tum
9	70	78.7	2000	4	AAC88152 Human K12
10	70	78.7	2000	6	ABL65927 Lung canc
11	70	78.7	2000	6	ABK84431 Human CDN
12	70	78.7	2000	12	ADJ74845 Marker ge
13	70	78.7	2000	13	ADR52746 Drug ther
14	70	78.7	2000	13	ADP24663 PRO polyp
15	70	78.7	2000	13	ADU06261 Novel bro
16	70	78.7	2000	14	ADW38555 Immunomod
17	70	78.7	2180	2	AAV32446 K12 cDNA
18	58.6	65.8	212	2	AAT26356 Human gen
19	54	60.7	401	10	ADF80286 Leukaemia

20	43.4	48.8	1131	12	ADQ24332	Adq24332 Human sof	
21	41.4	46.5	371	4	RAI87277	Aai87277 Human pol	
22	41.4	46.5	2066	12	ADK71117	Adk71117 Human CSN	
23	41.2	46.3	216215	10	ADP69167	Adf69167 Human MP5	
24	41	46.1	407	4	AAI89750	Aai89750 Human pol	
c	25	40.8	1155	6	ABL32605	Abi32605 Human imm	
26	40.6	45.6	1477	4	RAC84569	Aac84569 Rice magn	
c	27	40.4	45.4	1533	12	ADQ25148	Adq25148 Human sof
c	28	40.2	45.2	11745	6	ABK28331	Abk28331 DNA trans
c	29	39.6	44.5	1858	5	AAS68278	Aas68278 DNA encod
c	30	39.6	44.5	6207	4	AAS46777	Aas46777 Tumour su
c	31	39.4	44.3	238	6	ABL93031	Abi93031 Rat metas
c	32	39.4	44.3	238	12	ADN07822	Adn07822 Human mam
c	33	39.4	44.3	313	13	ACN58609	Acn58609 Cotton gy
c	34	39.4	44.3	990	10	ADRS5221	Adrs5221 Rat gene
c	35	39.4	44.3	6827	6	ABL33180	Abi33180 Human imm
c	36	39.4	44.3	6827	6	ABQ67019	Abq67019 Human ang
c	37	39.2	44.0	285	5	ADL42552	Adl42552 Human ova
c	38	39.2	44.0	379	4	AAL20476	Aal20476 Human bre
c	39	38.8	43.6	441	11	ACN92240	Acn92240 Breast ca
c	40	38.8	43.6	607	13	ACN58692	Acn58692 Cotton gy
c	41	38.8	43.6	781	4	AAL24400	Aal24400 Human bre
c	42	38.8	43.6	787	2	AAT29008	Aat29008 Parietari
c	43	38.8	43.6	11416	6	ABL32118	Abi32118 Human imm
c	44	38.8	43.6	11416	6	ABL70135	Abi70135 Chemicall
c	45	38.8	43.6	11416	6	AAS61063	Aas61063 Human gen

ALIGNMENTS

RESULT 1	
ID AAC89735	standard; cDNA; 90 BP.
XX	
AC AAC89735;	
XX	
DT 12-MAR-2001	(first entry)
XX	
DE Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 4.	
XX	
KW Human; cytostatic; immunomodulator; immunostimulant; vulnerary;	
KW anti-inflammatory; neuroprotective; antibacterial; gene therapy;	
KW gastrointestinal inflammation; immune system disorder; genetic disorder;	
KW cancer; autoimmune disorder; infection; wound healing; ss.	
XX	
OS Homo sapiens.	
XX	
PN WO200073324-A2.	
XX	
PD 07-DEC-2000.	
XX	
PF 01-JUN-2000; 2000MO-US015191.	
XX	
PR 01-JUN-1999; 99US-0137058P.	
XX	
PA (DIGI-) DIGITAL GENE TECHNOLOGIES INC.	
XX	
PI Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Haeel KW;	
XX	
DR WPI; 2001-061508/07.	
XX	
PT New polynucleotides and polypeptides, useful in gene therapy and in	
PT diagnosing a pathological condition, e.g. for modulating gene expression	
PT disorders.	
XX	
PS Claim 1; Page 85; 108pp; English.	
CC	
CC The present sequence is one of a number of isolated human polynucleotides	
CC which are useful in gene therapy, and for diagnosing a pathological	
CC condition or a susceptibility to it. In particular, the polynucleotides	
CC are useful for modulating gene expression in gastrointestinal	

CC inflammation. The polynucleotides are useful for chromosome
 CC identification, controlling gene expression through triple helix
 CC formation or antisense DNA or RNA, or identifying individuals from minute
 CC biological samples using DNA-based identification techniques. The
 CC polynucleotides can also be used as an alternative to restriction
 CC fragment length polymorphism (RFLP), by determining the actual base-by-
 CC base DNA sequences of selected portions of an individual's genome. The
 CC polynucleotides may also be used as molecular weight markers on Southern
 CC gels, as diagnostic probes for the presence of a specific mRNA, as a
 CC probe to subtract-out known sequences in the process of discovering novel
 CC polynucleotides, or as an antigen to elicit an immune response. The
 CC polypeptides are useful in diagnostic procedures to detect a disorder.
 CC The polynucleotides and polypeptides are useful for preventing, treating
 CC or ameliorating immune system disorders, genetic disorders, cancers, some
 CC autoimmune disorders, or infections. The polynucleotides and polypeptides
 CC are also useful for differentiating, proliferating or attracting cells,
 CC leading to the regeneration of tissues, especially in wounds or burns.
 CC The polypeptides and polynucleotides may also be used as a food additive
 CC or preservative
 XX
 SQ Sequence 90 BP; 46 A; 16 C; 12 G; 16 T; 0 U; 0 Other;

Query Match 100.0%; Score 89; DB 4; Length 90;
 Best Local Similarity 100.0%; Pred. No. 1.5e-10;
 Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 DB 1 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 QY 61 AACAAACCAAAAAAAAAAAAAAAAAAAAAA 89
 |||||
 DB 61 AACAAACCAAAAAAAAAAAAAAAAAAAAAA 89

RESULT 2
 AAC59772
 ID AAC59772 standard; cDNA; 1464 BP.

XX AAC59772;

XX 26-JAN-2001 (first entry)

XX Human secreted protein gene 35 SEQ ID NO:45.

XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
 KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
 KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
 KW fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
 KW hyperproliferative disorder; cancer; cardiovascular disorder; infection;
 KW cerebrovascular disorder; angiogenesis; nervous system disorder;
 KW ocular disorder; wound healing; skin aging; food additive; preservative;
 XX ss.

XX Homo sapiens.

XX WO200056751-A1.

XX 28-SEP-2000.

XX 09-MAR-2000; 2000WO-US006013.

XX 19-MAR-1999; 99US-0125360P.

XX 11-JUN-1999; 99US-0138626P.

XX 03-DEC-1999; 99US-0168662P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-579482/54.

XX P-PSDB; AAB34611.

PT Isolated nucleic acid molecule encoding a human secreted protein is used
 PT in preventing, treating or ameliorating a medical condition.
 XX
 PS Claim 1; Page 362; 419pp; English.

XX The polynucleotide sequences given in AAC59738 to AAC59787 encode the
 CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to
 CC AAB34686 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Example of activities include:
 CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;
 CC cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
 CC neuroprotective; antibacterial; virucide; fungicide; and
 CC ophthalmological. The polynucleotides and proteins can be used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 CC in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and
 CC cancers of the breast or liver, cardiovascular disorders, cerebrovascular
 CC disorders, angiogenesis, nervous system disorders, infections caused by
 CC bacteria, viruses and fungi and ocular disorders. The proteins can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The proteins can also be used as a
 CC food additive or preservative to increase or decrease storage
 CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used
 CC in the exemplification of the present invention

XX Sequence 1464 BP; 309 A; 483 C; 408 G; 264 T; 0 U; 0 Other;

Query Match 89.2%; Score 79.4; DB 3; Length 1464;
 Best Local Similarity 93.3%; Pred. No. 1.9e-08;
 Matches 83; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 DB 1071 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1130
 |||||
 QY 61 AACAAACCAAAAAAAAAAAAAAAAAAAAAA 89
 |||||
 DB 1131 AACAAACCGACTCACAAAAAAAAAAAAAAAAA 1159

RESULT 3
 AAC89734

ID AAC89734 standard; cDNA; 79 BP.

XX AAC89734;

XX 12-MAR-2001 (first entry)

XX Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 3.

XX Human; cytostatic; immunomodulator; immunostimulant; vulnerary;
 KW anti-inflammatory; neuroprotective; antibacterial; gene therapy;
 KW gastrointestinal inflammation; immune system disorder; genetic disorder;
 KW cancer; autoimmune disorder; infection; wound healing; ss.

XX Homo sapiens.

XX WO200073324-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015191.

XX 01-JUN-1999; 99US-0137058P.

XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.

PI Youakim A, Dubose RF, Sims JE, Pribyl TW, Hillbush BS, Hasel KW;
XX WPI; 2001-061508/07.
XX
XX New polynucleotides and polypeptides, useful in gene therapy and in
PT diagnosing a pathological condition, e.g. for modulating gene expression
PT in gastrointestinal inflammation, or for treating cancers or genetic
PT disorders.
XX
XX
XX Claim 1; Page 84; 108pp; English.
PS
XX The present sequence is one of a number of isolated human polynucleotides
CC which are useful in gene therapy, and for diagnosing a pathological
CC condition or a susceptibility to it. In particular, the polynucleotides
CC are useful for modulating gene expression in gastrointestinal
CC inflammation. The polynucleotides are useful for chromosome
CC identification, controlling gene expression through triple helix
CC formation or antisense DNA or RNA, or identifying individuals from minute
CC biological samples using DNA-based identification techniques. The
CC polynucleotides can also be used as an alternative to restriction
CC fragment length polymorphism (RFLP), by determining the actual base-by-
CC base DNA sequences of selected portions of an individual's genome. The
CC polynucleotides may also be used as molecular weight markers on Southern
CC gels, as diagnostic probes for the presence of a specific mRNA, as a
CC probe to subtract-out known sequences in the process of discovering novel
CC polynucleotides, or as an antigen to elicit an immune response. The
CC polypeptides are useful in diagnostic procedures to detect a disorder.
CC The polynucleotides and polypeptides are useful for preventing, treating
CC or ameliorating immune system disorders, genetic disorders, cancers, some
CC autoimmune disorders, or infections. The polynucleotides and polypeptides
CC are also useful for differentiating, proliferating or attracting cells,
CC leading to the regeneration of tissues, especially in wounds or burns.
CC The polypeptides and polynucleotides may also be used as a food additive
CC or preservative
XX
XX Sequence 79 BP; 31 A; 18 C; 12 G; 18 T; 0 U; 0 Other;
SQ

Query Match 79.8%; Score 71; DB 4; Length 79;
Best Local Similarity 93.7%; Pred. No. 1.3e-06;
Matches 74; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTTCCTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1 CGGGAATCCCTTCAGTTCCTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60

Qy 61 AACAAACCAAAAAA 79
Db 61 AACAAACCGACTCAAAA 79

RESULT 4
ADU14748
ID ADU14748 standard; DNA; 428 BP.
XX
XX ADU14748;
XX
XX 27-JAN-2005 (first entry)
XX
XX Solid tumour prognosis gene seqid 5187.
XX
XX cytostatic; gene therapy; expression profile; solid tumour;
KW peripheral blood mononuclear cell; PEMC; prognosis; ds.
XX
XX Unidentified.
XX
XX WO2004097052-A2.
XX
XX 11-NOV-2004.
XX
XX 29-APR-2004; 2004WO-US013587.
XX
XX 29-APR-2003; 2003US-0466067P.
PR
XX 23-JAN-2004; 2004US-0538246P.
PR

XX (AMHP) WYETH.
PA (STRA/) STRAHS A.
XX
XX Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
PI Immerman F, Dornier AU;
XX
XX WPI; 2004-804779/79.
XX
XX A method, useful for prognosing and treating solid tumor, comprises
PT comparing an expression profile of a gene expressed in peripheral blood
PT mononuclear cells to a reference expression profile of a gene.
XX
XX Disclosure; Page: 111pp; English.
XX
XX The invention describes a method comprising comparing an expression
CC profile of at least one gene in a peripheral blood sample of a patient to
CC at least one reference expression profile of the at least one gene, where
CC the patient has a solid tumour, and each of the gene is differentially
CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
CC of patients as compared to PBMCs of a second class of patients, where
CC both the first and second classes of patients have the solid tumour, and
CC each of the first and second classes is a subcluster formed by an
CC unsupervised clustering analysis of gene expression profiles in PBMCs of
CC a population of patients who have the solid tumour, and where the
CC majority of the first class of patients has a first clinical outcome, and
CC the majority of the second class of patients has a second clinical
CC outcome. Also described are: a system comprising (i) a memory or a
CC storage medium including data that represent an expression profile of at
CC least one gene in a peripheral blood sample of a patient who has a solid
CC tumour, (ii) at least another storage medium including data that
CC represent at least one reference expression profile of the gene, (iii) a
CC program capable of comparing the expression profile to the reference
CC expression profile, and (iv) a processor capable of executing the
CC program, where expression levels of the gene in peripheral blood
CC mononuclear cells of patients who have the solid tumour correlate with
CC clinical outcomes of the patients; and a nucleic acid or protein array
CC comprising concentrated probes for solid tumour prognosis genes, where
CC each of the solid tumour prognosis genes is differentially expressed in
CC PBMCs of a first class of patients as compared to PBMCs of a second class
CC of patients, where both the first and second classes of patients have a
CC solid tumour, and where the first class of patients has a first clinical
CC outcome, and the second class of patients has a second clinical outcome.
CC The method, system, and array are useful for prognosing and treating
CC solid tumors. This sequence represents a solid tumour prognosis gene of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 428 BP; 107 A; 108 C; 112 G; 81 T; 0 U; 20 Other;
SQ

Query Match 78.7%; Score 70; DB 13; Length 428;
Best Local Similarity 93.6%; Pred. No. 2.1e-06;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTTCCTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 100 CGGGAATCCCTTCAGTTCCTTGAAGAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 159

Qy 61 AACAAACCAAAAAA 78
Db 160 AACAAACCGACTCAAAA 177

RESULT 5
AEA36210/c
ID AEA36210 standard; DNA; 522 BP.
XX
XX AEA36210;
XX
XX 25-AUG-2005 (first entry)
XX
XX Human nucleic acid sequence #142.
DE

XX Screening; gene expression; colorectal tumor; colitis; Crohns disease;
KW irritable bowel syndrome; gastrointestinal disease; cytostatic;
KW gastrointestinal-gen.; antiinflammatory; ds.
XX Homo sapiens.
OS
XX WO2005054507-A2.
PN
XX 16-JUN-2005.
PD
XX
XX 03-DEC-2004; 2004WO-GB005078.
PF
XX
XX 04-DEC-2003; 2003GB-00028048.
PR
XX (UYSH-) UNIV SHEFFIELD.
PA
XX Corfe B, Chirakkal H;
XX
XX WPI; 2005-435407/44.
DR
XX
XX Screening for nucleic acid molecules exhibiting altered expression in
PT cells grown in the presence of butyrate, and detection of the nucleic
PT acid molecules or the encoded polypeptides in diagnosing colorectal
PT cancer.
PT
XX Disclosure; Page 186; 266pp; English.
PS
XX
XX The invention relates to a method of screening for nucleic acid molecules
CC that show altered expression in a first cell sample comprising comparing
CC the gene expression profile of the sample with that of a second reference
CC sample, where the first sample has been grown in the presence of butyrate
CC or a related carbon source from which butyrate is directly or indirectly
CC derived, but the reference sample has not. The invention also relates to
CC a method of detecting at least one nucleic acid molecule associated with
CC the initiation and/or progression of colorectal cancer in an animal,
CC comprising providing a biological sample comprising at least one cell to
CC be tested, contacting the sample with a ligand (preferably a hybridizing
CC nucleic acid molecule) which binds to at least one nucleic acid and
CC detecting the presence of at least one molecule in the sample, a method
CC of detecting at least one polypeptide associated with the initiation
CC and/or progression of colorectal cancer in an animal comprising providing
CC a biological sample comprising at least one cell to be tested, contacting
CC the sample with at least one ligand that specifically binds at least one
CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising
CC an amino acid sequence which varies by the addition, deletion or
CC substitution of at least one amino acid residue and detecting the
CC presence of the polypeptide in the sample, a method of screening for
CC agents that modulate the activity of at least one polypeptide encoded by
CC a gene associated with the initiation and/or progression of colorectal
CC cancer comprising forming a preparation comprising at least one
CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising
CC at least one addition, deletion or substitution and at least one agent to
CC be tested and determining the activity of the agent with respect to
CC activity of the polypeptide, and an antibody or its effective binding
CC portion identified by the method, for use as a pharmaceutical. The
CC methods are useful for screening for nucleic acid molecules that show
CC altered expression in a cell sample, and for detecting a nucleic acid and
CC a polypeptide respectively, that are associated with the initiation
CC and/or progression of colorectal cancer and are useful for detecting or
CC monitoring colorectal cancer, especially adenocarcinoma. The methods are
CC also useful for screening for agents that modulate the activity of at
CC least one polypeptide encoded by a gene associated with the initiation
CC and/or progression of cancer, where agents identified by the method are
CC useful for treating colorectal cancer. The methods could also be used to
CC detect or monitor other conditions such as colitis, Crohn's disease or
CC irritable bowel syndrome, as a screening tool for fiber consumption, as
CC an assay for colon microflora functionality or for early detection of pre
CC -cancerous growth. This sequence represents a human nucleic acid
CC identified by the screening method of the invention.
XX
SQ Sequence 522 BP; 99 A; 171 C; 170 G; 82 T; 0 U; 0 Other;

Query Match 78.7%; Score 70; DB 14; Length 522;
Best Local Similarity 93.6%; Pred. No. 2.1e-06;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
DB 82 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 23
QY 61 AACAAACCCAAAAA 78
DB 22 AACAAACCCGACTCACAA 5
RESULT 6
ADF81837
ID ADF81837 standard; DNA; 592 BP.
XX
AC ADF81837;
XX
DT 26-FEB-2004 (first entry)
XX
DE Leukaemia-related DNA sequence #2393.
XX
KW Cytostatic; Gene therapy; leukaemia; ss.
XX
OS Unidentified.
XX
PN WO2003039443-A2.
XX
PD 15-MAY-2003.
XX
PF 04-NOV-2002; 2002WO-BP012303.
XX
PR 05-NOV-2001; 2001EP-00126244.
PR 30-APR-2002; 2002EP-00009758.
XX
(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA (UULU-) UNIV LUDWIG MAXIMILIANS.
PA (HAFE/) HAFERLACH T.
PA (SCHO/) SCHOCH C.
PA (KERN/) KERN W.
XX
XX Haferlach T, Schoch C, Kern W, Kohlmann A, Schnittger S, Dugas M;
PI Eils R, Brors B, Mergenthaler S;
XX
XX WPI; 2003-505037/47.
XX
XX Determining the subtype of leukemia cells and whether a patient sample
PT contains leukemia cells or other cells, useful for treating leukemia,
PT comprises determining the expression profile of a group of markers in a
PT patient sample.
XX
XX Disclosure; SEQ ID NO 2393; 2938pp; English.
PS
XX The present invention relates to a method (M1) for determining the
CC subtype of leukemia cells and whether a patient sample contains
CC leukemia cells. The method comprises determining the expression profile
CC of a group of markers in a patient sample. The method is useful for
CC determining the presence of leukemia cells, its types or subtypes, and
CC for the preparation of a medicament for treating leukemia.
XX
SQ Sequence 592 BP; 133 A; 173 C; 178 G; 108 T; 0 U; 0 Other;
Query Match 78.7%; Score 70; DB 10; Length 592;
Best Local Similarity 93.6%; Pred. No. 2.1e-06;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
DB 434 CGGGAATCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 493
QY 61 AACAAACCCAAAAA 78
||||| | | | |

Db 494 AACAAACCGACTCAGAA 511

RESULT 7

ADQ86642

ID ADQ86642 standard; cDNA; 1985 BP.

XX AC

XX ADQ86642;

XX AC

XX 07-OCT-2004 (first entry)

XX DT

XX 07-OCT-2004 (first entry)

XX DE Human tumour-associated antigenic target (TAT) cDNA sequence #3516.

XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;

XX KW cancer; cell proliferative disorder; gene; ss.

XX KW

XX OS Homo sapiens.

XX OS

XX WO2004060270-A2.

XX FN

XX PD 22-JUL-2004.

XX PF 15-OCT-2003; 2003WO-US029126.

XX PF

XX 18-OCT-2002; 2002US-0418988P.

XX PR

XX (GETH) GENENTECH INC.

XX PA (WUTD/) WU T D.

XX PA (ZHOU/) ZHOU Y.

XX FI Wu TD, Zhou Y;

XX DR WPI; 2004-534300/51.

XX New nucleic acid molecule and encoded polypeptide, for diagnosing,

XX PT preventing or treating cell proliferative disorders such as cancer.

XX XX

XX Claim 1; SEQ ID NO 3516; 5504pp; English.

XX

XX The present invention describes an isolated tumour-associated antigenic

XX target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide

XX sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of

XX (a); (c) the complement of (a) or (b); (d) a sequence that has 80%

XX sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-

XX (c). Also described: (1) an expression vector comprising the above

XX nucleic acid; (2) a host cell comprising the above expression vector; (3)

XX a process for producing a polypeptide; (4) an isolated polypeptide

XX comprising: (a) an amino acid sequence encoded by any of the above

XX nucleotide sequences; (b) an amino acid sequence encoded by the full-

XX length coding region of the above nucleotide sequences; or (c) a sequence

XX having at least 80% identical to (a) or (b); (5) a chimeric polypeptide

XX comprising the above polypeptide fused to a heterologous polypeptide; (6)

XX an isolated antibody that binds to the above polypeptide; (7) a process

XX for producing the antibody; (8) an isolated oligopeptide that binds to

XX the above polypeptide; (9) a tumour-associated antigenic target (TAT)

XX binding organic molecule that binds to the above polypeptide; (10) a

XX composition of matter comprising the above (chimeric) polypeptide,

XX antibody, oligopeptide or TAT binding organic molecule, in combination

XX with a carrier; (11) an article of manufacture comprising a container and

XX the composition of matter contained within the container; (12) methods of

XX inhibiting the growth of a cell that expresses the above protein, where

XX the growth of the cell is at least in part dependent upon a growth

XX potentiating effect of the above protein; (13) a method of

XX therapeutically treating a mammal having a cancerous tumour comprising

XX cells that express the above protein; (14) a method of determining the

XX presence of a protein in a sample suspected of containing the protein

XX described above; (15) methods of diagnosing the presence of a tumour in a

XX mammal; (16) a method for treating or preventing a cell proliferative

XX disorder associated with increased expression or activity of the above

XX protein; and (17) a method of binding an antibody, oligopeptide or

XX organic molecule to a cell that expresses the protein described above.

XX The TAT sequences have cytostatic activities, and can be used in gene

XX therapy. The composition and methods are useful for diagnosing,

CC preventing or treating cancer. The composition is also used for preparing

CC a medicament for the therapeutic treatment or diagnostic detection of a

CC cell proliferative disorder or cancer. The present sequence represents a

CC human TAT cDNA sequence from the present invention.

XX

XX Sequence 1985 BP; 372 A; 649 C; 594 G; 370 T; 0 U; 0 Other;

Query Match 78.7%; Score 70; DB 12; Length 1985;

Best Local Similarity 93.6%; Pred. No. 2.1e-06;

Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 CGGGAATCCCTTCAGTTCTTTGAAAAAGCTTCATGACTCGATATATCTGAAATGAAGAA 60

Db 1636 CGGGAATCCCTTCAGTTCTTTGAAAAAGCTTCATGACTCGATATATCTGAAATGAAGAA 1695

Oy 61 AACAAACCAAAAAA 78

Db 1696 AACAAACCGACTCAGAA 1713

RESULT 8

ADQ85525

ID ADQ85525 standard; cDNA; 1985 BP.

XX AC

XX ADQ85525;

XX AC

XX 07-OCT-2004 (first entry)

XX DT

XX Human tumour-associated antigenic target (TAT) cDNA sequence #2339.

XX DE human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;

XX KW cancer; cell proliferative disorder; gene; ss.

XX KW

XX OS Homo sapiens.

XX OS

XX WO2004060270-A2.

XX FN

XX PD 22-JUL-2004.

XX PF 15-OCT-2003; 2003WO-US029126.

XX PF

XX 18-OCT-2002; 2002US-0418988P.

XX PR

XX (GETH) GENENTECH INC.

XX PA (WUTD/) WU T D.

XX PA (ZHOU/) ZHOU Y.

XX FI Wu TD, Zhou Y;

XX DR WPI; 2004-534300/51.

XX New nucleic acid molecule and encoded polypeptide, for diagnosing,

XX PT preventing or treating cell proliferative disorders such as cancer.

XX XX

XX Claim 1; SEQ ID NO 2339; 5504pp; English.

XX

XX The present invention describes an isolated tumour-associated antigenic

XX target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide

XX sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of

XX (a); (c) the complement of (a) or (b); (d) a sequence that has 80%

XX sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-

XX (c). Also described: (1) an expression vector comprising the above

XX nucleic acid; (2) a host cell comprising the above expression vector; (3)

XX a process for producing a polypeptide; (4) an isolated polypeptide

XX comprising: (a) an amino acid sequence encoded by any of the above

XX nucleotide sequences; (b) an amino acid sequence encoded by the full-

XX length coding region of the above nucleotide sequences; or (c) a sequence

XX having at least 80% identical to (a) or (b); (5) a chimeric polypeptide

XX comprising the above polypeptide fused to a heterologous polypeptide; (6)

XX an isolated antibody that binds to the above polypeptide; (7) a process

XX for producing the antibody; (8) an isolated oligopeptide that binds to

XX the above polypeptide; (9) a tumour-associated antigenic target (TAT)

XX binding organic molecule that binds to the above polypeptide; (10) a

XX composition of matter comprising the above (chimeric) polypeptide,

XX antibody, oligopeptide or TAT binding organic molecule, in combination

XX with a carrier; (11) an article of manufacture comprising a container and

XX the composition of matter contained within the container; (12) methods of

XX inhibiting the growth of a cell that expresses the above protein, where

XX the growth of the cell is at least in part dependent upon a growth

XX potentiating effect of the above protein; (13) a method of

XX therapeutically treating a mammal having a cancerous tumour comprising

XX cells that express the above protein; (14) a method of determining the

XX presence of a protein in a sample suspected of containing the protein

XX described above; (15) methods of diagnosing the presence of a tumour in a

XX mammal; (16) a method for treating or preventing a cell proliferative

XX disorder associated with increased expression or activity of the above

XX protein; and (17) a method of binding an antibody, oligopeptide or

XX organic molecule to a cell that expresses the protein described above.

XX The TAT sequences have cytostatic activities, and can be used in gene

XX therapy. The composition and methods are useful for diagnosing,

CC composition of matter comprising the above (chimeric) polypeptide.
 CC antibody, oligopeptide or TAT binding organic molecule, in combination
 CC with a carrier; (11) an article of manufacture comprising a container and
 CC the composition of matter contained within the container; (12) methods of
 CC inhibiting the growth of a cell that expresses the above protein, where
 CC the growth of the cell is at least in part dependent upon a growth
 CC potentiating effect of the above protein; (13) a method of
 CC therapeutically treating a mammal having a cancerous tumour comprising
 CC cells that express the above protein; (14) a method of determining the
 CC presence of a protein in a sample suspected of containing the protein
 CC described above; (15) methods of diagnosing the presence of a tumour in a
 CC mammal; (16) a method for treating or preventing a cell proliferative
 CC disorder associated with increased expression or activity of the above
 CC protein; and (17) a method of binding an antibody, oligopeptide or
 CC organic molecule to a cell that expresses the protein described above.
 CC The TAT sequences have cytostatic activities, and can be used in gene
 CC therapy. The composition and methods are useful for diagnosing,
 CC preventing or treating cancer. The composition is also used for preparing
 CC a medicament for the therapeutic treatment or diagnostic detection of a
 CC cell proliferative disorder or cancer. The present sequence represents a
 CC human TAT cDNA sequence from the present invention.

SQ Sequence 1985 BP; 372 A; 649 C; 594 G; 370 T; 0 U; 0 Other;

Query Match 78.7%; Score 70; DB 13; Length 1985;
 Best Local Similarity 93.6%; Pred. No. 2.1e-06;
 Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 Db 1636 CGGGAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1695

Qy 61 AACAAACCAAAAAA 78
 |||||

Db 1696 AACAAACCGACTCAGAA 1713

RESULT 9
 AAC88152

ID AAC88152 standard; cDNA; 2000 BP.

AC AAC88152;

DT 13-MAR-2001 (first entry)

DE Human K12 nucleotide sequence SEQ ID NO:3.

KW Human; CD7; K12; cognate ligand; cluster of differentiation; cancer;
 KW identification; inhibiting T cell proliferation; HIV; infection;
 KW activating natural killer cell proliferation; leukaemia; lymphoma;
 KW sepsis; graft versus host disease; autoimmune disease; arthritis;
 KW multiple sclerosis; rheumatoid arthritis; psoriatic arthritis; lupus;
 KW scleroderma; psoriasis; atopic dermatitis; type I diabetes mellitus;
 KW Hashimoto's thyroiditis; pernicious anaemia; Addison's disease; uveitis;
 KW myasthenia gravis; psoriasis; Guillain-Barre Syndrome; Grave's disease;
 KW systemic lupus erythematosus; dermatomyositis; asthma; eczema;
 KW atypical dermatitis; contact dermatitis; eczematous dermatitis;
 KW seborrheic dermatitis; rhinitis; ss.

OS Homo sapiens.

XX WO200073333-A2.

PN 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US014612.

XX 28-MAY-1999; 99US-0136450P.

XX (IMMUNEX CORP.

XX Lyman SD, Fanslow WC;

XX

DR WPI; 2001-061511/07.
 DR P-PSDB; AAB36658.

PT Stimulating intracellular signaling of CD7 comprises contacting a cell
 PT expressing CD7 with recombinant K12 protein, the cognate ligand of CD7,
 PT to inhibit T cell proliferation and/or activate natural killer cell
 PT proliferation.

XX Disclosure; Page 36-38; 42pp; English.

XX The present invention describes a method for stimulating (S) the
 CC intracellular signalling of CD (cluster of differentiation) 7 comprising
 CC contacting a cell that expresses CD7 with a recombinant K12 protein (I),
 CC the cognate ligand of CD7. (S) is useful for inhibiting T cell
 CC proliferation and/or activating NK (natural killer) cell proliferation of
 CC and/or inducing NK toxicity in a mammal which involves administration of
 CC K12 protein. It is also used for treating HIV-1 infection, cancer (T cell
 CC leukaemia, acute lymphocytic leukaemia, cutaneous T cell lymphoma),
 CC bacterial and viral infections, mediated by CD7. In the case of treating
 CC T cell leukaemia the soluble K12 protein is covalently attached to a
 CC toxin. A disease mediated by CD7 such as sepsis, graft versus host
 CC disease due to transplantation, autoimmune diseases, multiple sclerosis,
 CC arthritis, rheumatoid arthritis, psoriatic arthritis, scleroderma, lupus,
 CC psoriasis, pernicious anaemia, Addison's disease, myasthenia gravis,
 CC thyroiditis, atopic dermatitis, type I diabetes mellitus, Grave's disease,
 CC uveitis, psoriasis, Guillain-Barre Syndrome, Grave's disease, systemic
 CC lupus erythematosus and dermatomyositis, asthma, eczema, atypical
 CC dermatitis, contact dermatitis, other eczematous dermatides, seborrheic
 CC dermatitis, and rhinitis is also treated by administering a K12
 CC antagonist (neutralising antibody). The present sequence encodes the
 CC human K12 protein, which is given in the exemplification of the present
 CC invention

SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;

Query Match 78.7%; Score 70; DB 4; Length 2000;
 Best Local Similarity 93.6%; Pred. No. 2.1e-06;
 Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 Db 1645 CGGGAATCCCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
 |||||

Db 1705 AACAAACCGACTCAGAA 1722

RESULT 10

ABL65927

ID ABL65927 standard; DNA; 2000 BP.

XX ABL65927;

XX 15-MAY-2002 (first entry)

DT Lung cancer related gene sequence SEQ ID NO:4264.

DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.

XX Homo sapiens.

OS WO200194629-A2.

PN 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US010838.

XX 05-JUN-2000; 2000US-0209473P.

PR 05-JUN-2000; 2000US-0209531P.

CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
 CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful
 CC for screening an agent capable of modulating GCA preferably in an
 CC inflammation in a tissue; M4 is useful for detecting an inflammation
 CC (especially chronic) in a tissue, an allergic response in a subject,
 CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
 CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
 CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
 CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
 CC disease, ulcerative colitis, periodontal disease; also bacterial
 CC infection, viral infection, parasitic infection, protozoal infection,
 CC fungal infection and M5 is useful for treating one of the above
 CC conditions. The present sequence represents a gene differentially
 CC expressed in granulocytes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;
 Query Match 78.7%; Score 70; DB 6; Length 2000;
 Best Local Similarity 93.6%; Pred. No. 2.1e-06;
 Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 60
 DB 1645 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 1704
 QY 61 AACAAACCAAAAAA 78
 DB 1705 AACAAACCGACTCACAA 1722
 RESULT 12
 ADJ74845
 ID ADJ74845 standard; DNA; 2000 BP.
 XX
 AC ADJ74845;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Marker gene SEQ ID NO:97.
 XX
 KW bronchial asthma; chronic obstructive pulmonary disease;
 KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
 KW gene therapy; marker gene; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN EPI394274-A2.
 XX
 PD 03-MAR-2004.
 XX
 PF 04-AUG-2003; 2003EP-00254857.
 XX
 PR 06-AUG-2002; 2002JP-00229312.
 PR 20-MAR-2003; 2003JP-00077212.
 XX
 PA (GENO-) GENOX RES INC.
 XX
 PI Ontani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
 XX
 DR WPI; 2004-193155/19.
 XX
 PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
 PT comparing the expression level of a marker gene in a biological sample
 PT from a subject with the expression level of the gene in a sample from a
 PT healthy subject.
 XX
 PS Claim 1; SEQ ID NO 97; 241pp; English.
 XX
 CC The present invention describes a method of testing for bronchial asthma
 CC or chronic obstructive pulmonary disease. The method comprises

CC determining the expression level of a marker gene in a biological sample
 CC from a subject, comparing the expression level determined with the
 CC expression level of the marker gene in a biological sample from a healthy
 CC subject, and judging whether the subject has bronchial asthma or chronic
 CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
 CC genes (S1) whose expression levels increase when respiratory epithelial
 CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
 CC whose expression levels decrease when respiratory epithelial cells are
 CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
 CC testing for bronchial asthma or chronic obstructive pulmonary disease;
 CC (2) a kit for screening for a candidate compound for a therapeutic agent;
 CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
 CC an animal model for bronchial asthma or chronic obstructive pulmonary
 CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
 CC method for producing an animal model for bronchial asthma or chronic
 CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
 CC asthma or chronic obstructive pulmonary disease, comprising the compound,
 CC a marker gene or an antisense nucleic acid corresponding to a portion of
 CC the marker gene, a ribozyme, a polynucleotide that suppresses the
 CC expression of the gene through an RNAi effect or an antibody recognising
 CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
 CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
 CC probe has been immobilised to assay a marker gene. (I) has respiratory
 CC and antiasthmatic activities, and can be used in gene therapy. The method
 CC is useful for testing for or screening for a therapeutic agent for
 CC bronchial asthma or chronic obstructive pulmonary disease. The present
 CC sequence is used in the exemplification of the present invention.
 XX
 SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;
 Query Match 78.7%; Score 70; DB 12; Length 2000;
 Best Local Similarity 93.6%; Pred. No. 2.1e-06;
 Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 60
 DB 1645 CGGGAATCCCCCTTCAGTCTTTGAAAAGTTCATGACTCGAATATCTGAATGAAGAA 1704
 QY 61 AACAAACCAAAAAA 78
 DB 1705 AACAAACCGACTCACAA 1722
 RESULT 13
 ADR52746
 ID ADR52746 standard; DNA; 2000 BP.
 XX
 AC ADR52746;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Drug therapy altered expressed gene #97.
 XX
 KW drug activity monitoring; expression profile; gene expression;
 KW peripheral blood sample; peripheral blood mononuclear cell; drug therapy;
 KW CCI-779; immunosuppressant; rapamycin; mammalian target of rapamycin;
 KW mTOR; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2004072265-A2.
 XX
 PD 26-AUG-2004.
 XX
 PF 11-FEB-2004; 2004WO-US004118.
 XX
 PR 11-FEB-2003; 2003US-0446133P.
 PR 03-APR-2003; 2003US-0459782P.
 PR 23-JAN-2004; 2004US-0538246P.
 XX
 PA (AMHP) WYETH.
 PA (BURC/) BURCZYNSKI M.
 PA (TWIN/) TWINE N.

PA (DORN/) DORNER A J.
PA (TREP/) TREPICCHIO W L.
XX Burczynski M, Twine N, Dorner AJ, Trepicchio WL;
XX WPI; 2004-642301/62.
XX Monitoring drug activities in vivo comprises comparing an expression
XX profile of a gene in a peripheral blood sample of a patient before and
XX after drug therapy.
XX Disclosure; SEQ ID NO 97; 136pp; English.
XX The invention relates to a method of monitoring drug activities in vivo
XX by comparing an expression profile of at least one gene in a peripheral
XX blood sample of a patient to a reference expression profile of the at
XX least one gene, where the at least one gene is differentially expressed
XX in peripheral blood mononuclear cells (PBMCs) of patients who have a non-
XX blood disease and are subjected to a drug therapy as compared to PBMCs
XX isolated from the patient before the drug therapy, and where the patient
XX has the non-blood disease and is being treated by the drug therapy. The
XX method, kit, and nucleic acid array are useful for monitoring drug
XX activities in vivo. The drug is especially CCI-779, an ester analogue of
XX the immunosuppressant rapamycin which is a potent inhibitor of the
XX mammalian target of rapamycin (mTOR). This sequence represents a gene
XX expressed in PBMC altered by the drug therapy. (Note: this sequence does
XX no form part of the printed specification but was obtained in electronic
XX format from WIPO at ftp.wipo.int/pub/published_pat_sequences/).

XX Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;
XX Query Match 78.7%; Score 70; DB 13; Length 2000;
XX Best Local Similarity 93.6%; Pred. No. 2.1e-06;
XX Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Db 1705 AACAAACCGACTCACAAA 1722

RESULT 14
ADP24663
ID ADP24663 standard; cDNA; 2000 BP.
XX ADP24663;
XX 18-NOV-2004 (first entry)
XX PRO polypeptide encoding cDNA SEQ ID NO:1841.
XX ss: gene; PRO; antiinflammatory; antiarthritic; antirheumatic;
XX immunosuppressive; osteopathic; antidiabetic; dermatological;
XX antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;
XX gene therapy; immune system.
XX Unidentified.
XX WO2004041170-A2.
XX 21-MAY-2004.
XX 30-OCT-2003; 2003WO-US034312.
XX 01-NOV-2002; 2002US-0423394P.
XX (GETH) GENENTECH INC.
XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
XX Wu TD;

XX WPI; 2004-419628/39.
XX P-PSDB; ADP24664.
XX New PRO polypeptides and polynucleotides, useful for treating e.g.
XX erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
XX renal disease, or demyelinating diseases of the central or peripheral
XX nervous system.
XX Claim 1; SEQ ID NO 1841; 2940pp; English.
XX The invention relates to a novel isolated nucleic acid and the PRO
XX polypeptide encoded by it. A protein of the invention has
XX antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
XX osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
XX antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
XX of the invention may have a use in gene therapy. The PRO polypeptide, its
XX agonist, antagonist, or antibody that specifically binds to the
XX polypeptide is useful for treating an immune related disorder such as
XX systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
XX juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
XX idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
XX vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
XX thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
XX disease, a demyelinating disease of the central or peripheral nervous
XX system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
XX a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
XX disease, infectious or autoimmune chronic active hepatitis, primary
XX biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
XX inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
XX disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
XX disease, asthma, allergic rhinitis, atopic dermatitis, food
XX hypersensitivity, urticaria, an immunologic disease of the lung,
XX eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
XX pneumonitis, a transplantation associated disease, graft rejection or
XX graft-versus-host disease. The present sequence encodes a PRO protein of
XX the invention.

XX Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;
XX Query Match 78.7%; Score 70; DB 13; Length 2000;
XX Best Local Similarity 93.6%; Pred. No. 2.1e-06;
XX Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGGAATCCCCCTTCAGTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCAAAAAA 78
Db 1705 AACAAACCGACTCACAAA 1722

RESULT 15
ADU06261
ID ADU06261 standard; DNA; 2000 BP.
XX ADU06261;
XX 27-JAN-2005 (first entry)
XX Novel bronchial cancer-associated human gene SeqID485.
XX bronchial cancer; cytostatic; tumour-associated protein;
XX cancer detection; metastasis; tumour; gene; ds; human.
XX Homo sapiens.
XX DE10316701-A1.
XX 04-NOV-2004.

PF 09-APR-2003; 2003DE-01016701.
XX
PR 09-APR-2003; 2003DE-01016701.
XX
PA (HINZ/) HINZMANN B.
PA (HERM/) HERMANN K.
PA (CAST/) HEIDEN CASTANOS-VELEZ E.
XX
PI Mennerich D, Bruemendorf T, Heiden E, Hermann K, Kinnemann H;
PI Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pilarsky C;
XX
DR WPI; 2004-786403/78.
DR P-PSDB; ADU06748.
XX
PT New nucleic acid, and derived proteins, useful for diagnosis of bronchial
PT cancer and in screening for therapeutic and diagnostic agents.
XX
PS Claim 1; SEQ ID NO 485; 1381pp; German.
XX
CC This invention relates to a novel isolated nucleic acid associated with
CC bronchial cancer comprising 489 defined sequences given in the
CC specification. The invention may be useful for the production of
CC compounds with a cytostatic activity through the inhibition of expression
CC or activity of tumour-associated proteins. The novel DNA sequences and
CC the proteins/peptides encoded by them are used for detecting bronchial
CC cancer or determining the risk of developing it and to screen for
CC specific binding partners of the DNA or protein sequences, where the
CC binding partners are potentially useful as agents for treating or
CC diagnosing bronchial cancer. The DNA or protein sequences can also be
CC used for prognosis, detection of metastases and for secondary treatment
CC (of tumours that have been stabilised or are no longer detectable).
CC Detecting abnormal expression of the DNA sequences provides early
CC diagnosis of bronchial cancers. The present sequence is that of a novel
CC bronchial cancer-associated human gene sequence of the invention.
XX
SQ Sequence 2000 BP; 367 A; 659 C; 602 G; 372 T; 0 U; 0 Other;
Query Match 78.7%; Score 70; DB 13; Length 2000;
Best Local Similarity 93.6%; Pred. No. 2.1e-06;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CGGGAATCCCCCTTCACTTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGGAATCCCCCTTCACTTTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704
Qy 61 AACAAACCAAAAAAAA 78
Db 1705 AACAAACCGACTCACAAA 1722
Search completed: December 7, 2005, 00:03:01
Job time : 140.915 secs

;; PRIOR APPLICATION NUMBER: JP 2002-229312
;; PRIOR FILING DATE: 2002-08-06
;; NUMBER OF SEQ ID NOS: 2086
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 97
;; LENGTH: 2000
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-631-467-97

Query Match 78.7%; Score 70; DB 9; Length 2000;
Best Local Similarity 93.6%; Pred. No. 1.6e-07;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1645 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1704

Qy 61 AACAAACCAAAAAAAA 78
Db 1705 AACAAACCGACTCACA 1722

RESULT 6

US-10-080-522-2
;; Sequence 2, Application US/10080522
;; Publication No. US20030096326A1
;; GENERAL INFORMATION:
;; APPLICANT: KAUFMAN, RUSSEL E.
;; SLENTZ-KESLER, KIMBERLY
;; TITLE OF INVENTION: GENE PRODUCT OVER EXPRESSED IN CANCER
CELLS

;; NUMBER OF SEQUENCES: 2
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: NIXON & VANDERHYE P.C.
;; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
;; CITY: ARLINGTON
;; STATE: VIRGINIA
;; COUNTRY: U.S.A.
;; ZIP: 22201-4714

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/10/080,522
;; FILING DATE: 25-Feb-2002
;; CLASSIFICATION: <Unknown>

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 09/539,774
;; FILING DATE: 31-MAR-2000
;; APPLICATION NUMBER: US 09/210,474
;; FILING DATE: 14-DEC-1998
;; APPLICATION NUMBER: US 08/755,559
;; FILING DATE: 22-NOV-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: WILSON, MARY J.
;; REGISTRATION NUMBER: 32,955
;; REFERENCE/POCKET NUMBER: 1579-645
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 816-4000
;; TELEFAX: (703) 816-4100

;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2180 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (Genomic)

;; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-080-522-2

Query Match 78.7%; Score 70; DB 5; Length 2180;
Best Local Similarity 93.6%; Pred. No. 1.7e-07;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 1841 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 1900

Qy 61 AACAAACCAAAAAAAA 78
Db 1901 AACAAACCGACTCACA 1918

RESULT 7

US-10-723-860-7152
;; Sequence 7152, Application US/10723860
;; Publication No. US20040253606A1
;; GENERAL INFORMATION:
;; APPLICANT: Aziz, Natasha
;; APPLICANT: Ginsburg, Wendy M.
;; APPLICANT: Zlotnik, Albert
;; TITLE OF INVENTION: Methods for Diagnosis of Soft Tissue Sarcoma, Compositions &
;; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
;; FILE REFERENCE: 05882.0193.NPUS01
;; CURRENT APPLICATION NUMBER: US/10/723,860
;; PRIOR FILING DATE: 2003-11-26
;; PRIOR APPLICATION NUMBER: 60/429,739
;; PRIOR FILING DATE: 2002-11-26
;; NUMBER OF SEQ ID NOS: 8393
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 7152

;; LENGTH: 1131
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (12)..(32)
;; OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-7152

Query Match 48.8%; Score 43.4; DB 8; Length 1131;
Best Local Similarity 72.7%; Pred. No. 0.46;
Matches 56; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 13 TTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAACCAACCAAAA 72
Db 1000 TTAACACTTACAATAATCCAGACTGTCATATTTAAAAAGAAAAAAGAAAAA 1059

Qy 73 AAAAAAAGAAAAA 89
Db 1060 AAAAAAAGAAAAA 1076

RESULT 8

US-10-437-963-70672/c
;; Sequence 70672, Application US/10437963
;; Publication No. US20040123343A1
;; GENERAL INFORMATION:

;; APPLICANT: La Rosa, Thomas J.
;; APPLICANT: Kovalic, David K.
;; APPLICANT: Zhou, Yihua
;; APPLICANT: Cao, Yongwei
;; APPLICANT: Wu, Wei
;; APPLICANT: Boukharov, Andrey A.
;; APPLICANT: Barbazuk, Brad
;; APPLICANT: Li, Ping
;; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
;; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

;; FILE REFERENCE: 38-21(53221)B
;; CURRENT APPLICATION NUMBER: US/10/437,963
;; CURRENT FILING DATE: 2003-05-14
;; NUMBER OF SEQ ID NOS: 204966
;; SEQ ID NO 70672

```
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_71220C.1
US-10-437-963-70672

Query Match      48.5%; Score 43.2; DB 7; Length 433;
Best Local Similarity 68.2%; Pred. No. 0.42; Mismatches 0; Indels 0; Gaps 0;
Matches 60; Conservative 0;

Qy  2 GGGAAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGGAAGAAA 61
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  93 GGGAAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGGAAGAAA 34
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  62 ACAAACCAAAAAAAGAAAAAGAAAAA 89
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  33 AAAAAAAGAAAAAAGAAAAAAGAAAAA 6

RESULT 9
US-10-312-841-2/c
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (379615)
US-10-312-841-2

Query Match      47.9%; Score 42.6; DB 6; Length 3673778;
Best Local Similarity 78.5%; Pred. No. 4; Mismatches 0; Indels 0; Gaps 0;
Matches 51; Conservative 0;

Qy  25 AAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAACCAACCAAAAAAAGAAAAAAG 84
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  775209 AAAAAATTTAAAACTTCATATTTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 775150
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  85 AAAAA 89
    |||||
Db  775149 AAAAA 775145

RESULT 10
US-10-425-115-123809/c
; Sequence 123809, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 123809
; LENGTH: 397
; TYPE: DNA

; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_44396C.1
US-10-425-115-123809

Query Match      47.2%; Score 42; DB 8; Length 397;
Best Local Similarity 69.5%; Pred. No. 0.82; Mismatches 25; Indels 0; Gaps 0;
Matches 57; Conservative 0;

Qy  8 CCCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGGAAGAAAAACAAAC 67
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  139 CCCCATTAAGTTATTGAAGCCGTGGCTTCCTTGGGTTTCCCAAAAAAAGAAAAAAGAAA 80
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  68 CAAAAAAGAAAAAAGAAAAAAGAAA 89
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  79 AAAAAAAGAAAAAAGAAAAAAGAAA 58
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
US-10-425-115-158614/c
; Sequence 158614, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 158614
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_76234C.1
US-10-425-115-158614

Query Match      46.7%; Score 41.6; DB 8; Length 499;
Best Local Similarity 67.0%; Pred. No. 1.1; Mismatches 29; Indels 0; Gaps 0;
Matches 59; Conservative 0;

Qy  2 GGGAAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGGAAGAAA 61
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  133 GGGGGGCCCCCCCCAAGGAGTTTTGAAAAAACCACATGTTCCGAAAAAAGAAAAAAGAAA 74
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy  62 ACAAACCAAAAAAAGAAAAAAGAAAAA 89
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  73 AAAAAAAGAAAAAAGAAAAAAGAAA 46
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
US-10-425-115-31392/c
; Sequence 31392, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 31392
; LENGTH: 670
; TYPE: DNA
; ORGANISM: Zea mays
```

```
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_128637C.1
US-10-425-115-31392

Query Match      46.1%; Score 41; DB 8; Length 670;
Best Local Similarity 66.3%; Pred. No. 1.6;
Matches 59; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
Db 154 CGGGGCCCCCTTTTGTGGGAATTTGGTGTGTTTGTCTCAAAAAA 95

Qy 61 AACAAACCAAAAAAAGAAAAA 89
Db 94 AAAAAAAGAAAAAAGAAAAA 66

RESULT 13
US-10-425-115-100475/c
; Sequence 100475, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 100475
; LENGTH: 252
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(252)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_23146C.1
US-10-425-115-100475

Query Match      45.8%; Score 40.8; DB 8; Length 252;
Best Local Similarity 65.1%; Pred. No. 1.5;
Matches 54; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 7 TCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAACAAA 66
Db 105 TCCTCCCTATTATAAGAAATGTTTCATTGTTTAAANNNGAAAAAAGAAAAA 46

Qy 67 CCAAAAAAAGAAAAAAGAAAAA 89
Db 45 AAAAAAAGAAAAAAGAAAAA 23

RESULT 14
US-10-311-455-578/c
; Sequence 578, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: FIEFENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013,1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
```

```
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 578
; LENGTH: 11155
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-578

Query Match      45.8%; Score 40.8; DB 6; Length 11155;
Best Local Similarity 67.9%; Pred. No. 3.3;
Matches 57; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 6 ATCCCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAAACAA 65
Db 1913 ATAATCCATATTTATTACAAAAAATCCGAAACTTCAATATTTAAAAAAGAAAAAACA 1854

Qy 66 ACCAAAAAAGAAAAAAGAAAAA 89
Db 1853 ACAATAAAAAAAGAAAAAACA 1830

RESULT 15
US-10-437-963-15885/c
; Sequence 15885, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 15885
; LENGTH: 496
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(496)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_21687C.1
US-10-437-963-15885

Query Match      45.6%; Score 40.6; DB 7; Length 496;
Best Local Similarity 65.9%; Pred. No. 1.9;
Matches 58; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 2 GGGAAATCCCTTCAGTTCCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 61
Db 127 GGGGGCCCCCTTCGGGCTTTGNAAAAAAAGAAAAAAGAAAAAAGAAAAA 68

Qy 62 ACAAACCAAAAAAAGAAAAAAGAAAAA 89
Db 67 AAAAAAAGAAAAAAGAAAAAAGAAAAA 40

Search completed: December 6, 2005, 22:57:04
Job time : 226.613 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 1015.87 Seconds
(without alignments)
4099.008 Million cell updates/sec

Title: US-09-980-046B-4
Perfect score: 89
Sequence: 1 cgggaatcccccttcagtc.....aaaaaaaaaaaaaagaaaaa 89

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_hc: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_est7: *
9: gb_gss1: *
10: gb_gss2: *
11: gb_gss3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	84.2	94.6	693	CB048392	CB048392 NISC_gj04
C 2	82.8	93.0	519	AW014948	AW014948 UI-H-B10-
C 3	79.4	89.2	338	CK903980	CK903980 ie69h03.y
C 4	79.4	89.2	633	BQ182988	BQ182988 UI-H-ED1-
C 5	79.4	89.2	684	CB241372	CB241372 UI-CF-FN0
C 6	79.4	89.2	691	CD364877	CD364877 UI-H-F72-
C 7	75.4	84.7	313	BM053069	BM053069 ie69h03.y
C 8	73	82.0	292	BF431778	BF431778 nab50c04.
C 9	73	82.0	310	AA903087	AA903087 ok46d04.s
C 10	73	82.0	339	CK903979	CK903979 ie69h03.x
C 11	73	82.0	475	A1927624	A1927624 wo85a02.x
C 12	72	80.9	716	BF002331	BF002331 7h22b08.x
C 13	71.6	80.4	445	A1480347	A1480347 tm51c09.x
C 14	71.4	80.2	687	BU689727	BU689727 UI-CF-FN0
C 15	70	78.7	109	AA826896	AA826896 nr86c10.s
C 16	70	78.7	180	AW874230	AW874230 hg86h05.x
C 17	70	78.7	272	BF002304	BF002304 7h02g04.x
C 18	70	78.7	332	A1621135	A1621135 tu49d01.x
C 19	70	78.7	376	A1281889	A1281889 qt8aall.x
C 20	70	78.7	417	AW002382	AW002382 w61c03.x
C 21	70	78.7	418	AW594044	AW594044 hg35d12.x
C 22	70	78.7	420	AW195936	AW195936 xl85a06.x

C 23	70	78.7	429	1	AW207187	AW207187 UI-H-B11-
C 24	70	78.7	432	1	AW057702	AW057702 wx02c07.x
C 25	70	78.7	440	1	A1218726	A1218726 oo07a10.x
C 26	70	78.7	442	1	A1652267	A1652267 wb28d06.x
C 27	70	78.7	445	1	AW058447	AW058447 wx21a06.x
C 28	70	78.7	453	1	A1805092	A1805092 tu33g10.x
C 29	70	78.7	479	5	BX108872	BX108872 BX108872
C 30	70	78.7	522	2	BF939675	BF939675 nac79g07.
C 31	70	78.7	1633	4	CR595956	CR595956 full-leng
C 32	69	77.5	429	1	AW081554	AW081554 xc42g10.x
C 33	69	77.5	456	1	A1979001	A1979001 wr69a10.x
C 34	68.4	76.9	408	2	BE677472	BE677472 7d84g07.x
C 35	68.4	76.9	413	1	A1433486	A1433486 ti53b01.x
C 36	68.4	76.9	413	1	A1433494	A1433494 ti53c03.x
C 37	68.4	76.9	1650	4	CR594943	CR594943 full-leng
C 38	67.6	76.0	322	5	BX378127	BX378127 BX378127
C 39	66.8	75.1	157	1	A1277084	A1277084 ql40h08.x
C 40	66.8	75.1	350	1	AW974325	AW974325 EST386428
C 41	64.2	72.1	158	1	A1886304	A1886304 wm67a04.x
C 42	62.2	69.9	799	5	BX380143	BX380143 BX380143
C 43	61.6	69.2	1064	5	BX335528	BX335528 BX335528
C 44	58.6	65.8	209	5	C01591	C01591 HUMGS000859
C 45	58	65.2	1700	4	CR614987	CR614987 full-leng

ALIGNMENTS

RESULT 1
CB048392/c
LOCUS
DEFINITION NISC_gj04e10.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:3270643
3', mRNA sequence.
ACCESSION CB048392
VERSION CB048392.1 GI:27786679
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 693)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb-re@mail.nih.gov
CDNA Library Preparation:
CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plate: L1AM8006 row: J column: 20
Seq primer: -21M13 forward primer (ABI).
Location/Qualifiers
1. .693
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3270643"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Pr28"
/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI_CGAP_Pr22 was prepared, and ss
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool

of 5,000 clones made from the same library (cloneIDs 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo.

ORIGIN

Query Match 94.6%; Score 84.2; DB 6; Length 693;
Best Local Similarity 96.6%; Pred. No. 7.4e-09;
Matches 86; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGGAATCCCTTCAGTCTTTGAAAAAGTTCATGCTCGAATATCTGAATGAAGAA 60
|||||
Db 92 CGGAATCCCTTCAGTCTTTGAAAAAGTTCATGCTCGAATATCTGAATGAAGAA 33
|||||

Qy 61 AACAAACCAAAAAAAGAAAAAAGAAAAA 89
|||||
Db 32 AACAAACCGAAAAAAGAAAAAAGAAAAA 4
|||||

RESULT 2

AW014948/c
LOCUS AW014948 519 bp mRNA linear EST 10-SEP-1999
DEFINITION UI-H-BIO-aag-h-07-0-UI.s1 NCI_CGAP_Sub1 Homo sapiens cDNA clone
IMAGE:2709349 3', mRNA sequence.
ACCESSION AW014948
VERSION AW014948.1 GI:5863705
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

1 (bases 1 to 519)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL

Unpublished (1997)

Contact: Robert Straubeberg, Ph.D.

Email: csapbs-remail.nih.gov

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html The following repetitive

elements were found in this cDNA sequence: 271-336,

>GC richLow complexity

Seq primer: M13 Forward

POLYA=Yes.

FEATURES

source

Location/Qualifiers

1..519

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2709349"

/lab_host="NCI_CGAP_Sub1"

/note="Vector: p773D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; The

NCI_CGAP_Sub1 library is a subtracted library derived from

BI_BI Constitutes a mixture of 21 normalized or

subtracted NCI_CGAP libraries: NCI_CGAP_C04,

NCI_CGAP_Pr28, NCI_CGAP_Pr28, NCI_CGAP_C010,

NCI_CGAP_C016, NCI_CGAP_Kid3, NCI_CGAP_Kid12,

NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2,

NCI_CGAP_Br2, NCI_CGAP_C08, NCI_CGAP_CEL1, NCI_CGAP_Le12,

NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_Lu24,

NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6, NCI_CGAP_Brn25.

These 21 libraries were pooled and a single-stranded DNA

preparation of the resulting mixture was used as a tracer

in a subtractive hybridization with a driver whose

composition is detailed below: NCI_CGAP_Kid3 pool 1 LLAM

3334-3337, 3682-3683, 3798-3803 (IMAGE CloneIDs

ORIGIN

Query Match 93.0%; Score 82.8; DB 1; Length 519;
Best Local Similarity 97.7%; Pred. No. 1.5e-08;
Matches 84; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAATCCCTTCAGTCTTTGAAAAAGTTCATGCTCGAATATCTGAATGAAGAA 60
|||||

Db 86 CGGAATCCCTTCAGTCTTTGAAAAAGTTCATGCTCGAATATCTGAATGAAGAA 27
|||||

Qy 61 AACAAACCAAAAAAAGAAAAAAGAAAAA 86
|||||

Db 26 AACAAACCGAAAAAAGAAAAAAGAAAAA 1
|||||

RESULT 3

CK903980

LOCUS

DEFINITION

CK903980

ACCESSION

CK903980

VERSION

CK903980.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 338)

Melton D., Meadows A., Clifton S., Hillier L., Marra M., Pape D.,

Wyllie T., Martin J., Blistain A., Schmitt A., Theising B.,

Ritter E., Ronko I., Bennett J., Cardenas M., Gibbons M.,

McCann R., Cole R., Tsagareishvili R., Williams T., Jackson Y. and

Bowers Y.

WashU-Harvard Pancreas EST Project

Unpublished (2000)

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

This read is a 5' RESEQUENCE of a previously sequenced pancreas

clone

This read has been verified (found to hit its original self in the

correct orientation)

Putative full length read

vector to vector length is

Seq primer: -40RP from Gibco.

Location/Qualifiers

1..338

1322376-1323911, 1456008-1456775, 1500552-1502855)
NCI_CGAP_Kid3 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778
(IMAGE CloneIDs 1323912-1325831, 1471368-1472503,
1492104-1493255) NCI_CGAP_Lu5 pool 1 LLAM 3575-3582,
3851-3854 (IMAGE CloneIDs 1414920-1417991,
1520904-1522439) NCI_CGAP_GC4 pool 1 LLAM 3164-3167,
3716-3720, 3733-3735 (IMAGE CloneIDs 1257096-1258631,
1469064-1470983, 1475592-1476743) NCI_CGAP_Pr22 pool 1
LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE CloneIDs
985608-986759, 1101192-1101959, 1217928-1220615)
NCI_CGAP_C010 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE
CloneIDs 1057416-1061255, 1144584-1145351) The resulting
subtracted library contained 530,000 recombinants.
Subtraction was performed as previously described
[Bonaldo, Lennon & Soares (1996): Normalization and
Subtraction: Two Approaches To Facilitate Gene Discovery.
Genome Research 6, 791-806.
TAG_TISSUE=kidney
TAG_LIB=NCI_CGAP_Kid3
TAG_SEQ=ATTC

CK903980 338 bp mRNA linear EST 11-MAR-2004
te69h03.y5 Melton Normalized Human Islet 4 M4-HIS 1 Homo sapiens
cDNA clone IMAGE:5672308 5', mRNA sequence.

CK903980

ACCESSION

CK903980

VERSION

CK903980.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 338)

Melton D., Meadows A., Clifton S., Hillier L., Marra M., Pape D.,

Wyllie T., Martin J., Blistain A., Schmitt A., Theising B.,

Ritter E., Ronko I., Bennett J., Cardenas M., Gibbons M.,

McCann R., Cole R., Tsagareishvili R., Williams T., Jackson Y. and

Bowers Y.

WashU-Harvard Pancreas EST Project

Unpublished (2000)

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

This read is a 5' RESEQUENCE of a previously sequenced pancreas

clone

This read has been verified (found to hit its original self in the

correct orientation)

Putative full length read

vector to vector length is

Seq primer: -40RP from Gibco.

Location/Qualifiers

1..338


```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5672308"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="Mclon Normalized Human Islet 4 N4-HIS 1"
/notes="Organ: Pancreas; Vector: pSPOR1; Site 1: Not 1; Site 2: Sal 1; Starting library constructed using SuperScript Plasmid Library kit (Life Technologies). cDNA made by oligo-dT priming. Size-selected by column fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an Ecot of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

```

ORIGIN

```

Query Match      89.2%; Score 79.4; DB 7; Length 338;
Best Local Similarity 93.3%; Pred. No. 8.8e-08;
Matches 83; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
    |||||
Db 247 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 306
    |||||

QY 61 AACAAACCAAAAAAAGAAAAAAGAAAAA 89
    |||||
Db 307 AACAAACCGACTCACAAAAAAGAAAAA 335
    |||||

```

RESULT 4

```

BQ182988/c
LOCUS      BQ182988      633 bp      mRNA      linear      EST 15-JUL-2003
DEFINITION UI-H-ED1-ayj-1-19-0-UI.s2 NCI CGAP ED1 Homo sapiens cDNA clone
IMAGE:5839650 3', mRNA sequence.
ACCESSION  BQ182988
VERSION     BQ182988.1 GI:20358538
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1 (bases 1 to 633)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov

```

```

Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
The following repetitive elements were found in this cDNA
sequence: 276-341, >GC rich#Low_complexity
Seq primer: M13 FORWARD
POLYA=Yes.

```

FEATURES

```

source      Location/Qualifiers
1..633
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"

```

```

/clone="IMAGE:5839650"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP ED1"
/notes="Organ: Left Pubic Bone; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP ED1 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma cell line C55. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCTCAAGGCT.
TAG_TISSUE=chondrosarcoma
TAG_LIB=UI-H-ED1
TAG_SEQ=CGTCAAGGCT"

```

ORIGIN

```

Query Match      89.2%; Score 79.4; DB 3; Length 633;
Best Local Similarity 93.3%; Pred. No. 8.5e-08;
Matches 83; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
    |||||
Db 91 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 32
    |||||

QY 61 AACAAACCAAAAAAAGAAAAAAGAAAAA 89
    |||||
Db 31 AACAAACCGACTCACAAAAAAGAAAAA 3
    |||||

```

RESULT 5

```

CB241372/c
LOCUS      CB241372      684 bp      mRNA      linear      EST 12-FEB-2003
DEFINITION UI-CF-FNO-afy-h-23-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone
IMAGE:5839650 3', mRNA sequence.
ACCESSION  CB241372
VERSION     CB241372.1 GI:28363016
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1 (bases 1 to 684)
            Bonaldo,M.F., Lennon,G. and Soares,M.B.
            Normalization and subtraction: two approaches to facilitate gene
            discovery
            Genome Res. 6 (9), 791-806 (1996)
            8889548
            Contact: McCray, PB
            McCray Lab
            University of Iowa
            2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
            Tel: 319 356 4866
            Fax: 319 356 7171
            Email: paul-mccray@uiowa.edu

```

```

Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.reagen.com) or from Open Biosystems
(www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 277-342, >GC rich#Low_complexity
Seq primer: M13 FORWARD

```

```

POLYA=Yes.
FEATURES
  source      Location/Qualifiers
    1..684
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="UI-CF-FNO-afy-h-23-0-UI"
      /tissue_type="Human Lung Epithelial cells"
      /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
      /clone_lib="UI-CF-FNO"
      /notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
      modified polylinker; Site 1: Ecor I; Site 2: Not I;
      UI-CF-FNO is a subtracted cDNA library derived from two
      normalized Human lung epithelial cell libraries (EN1 and
      DUL) the library was subtracted according to according to
      Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
      1996. For additional information, contact:
      bento-soares@uiowa.edu
      TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
      6hr to LPS 24h
      TAG_LIB=UI-CF-FNO
      TAG_SEQ=CTGCTCAGGT"

ORIGIN
  Query Match      89.2%; Score 79.4; DB 6; Length 684;
  Best Local Similarity 93.3%; Pred. No. 8.5e-08;
  Matches 83; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGGGAATCCCCCTTCAGTCTTTTGAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
    |||||||
Db 92 CGGGAATCCCCCTTCAGTCTTTTGAAGTTCATGACTCGAATATCTGAAATGAAGAA 33
    |||||||

QY 61 AACAAACCAAAAAAAAAAAAAAAAAAGAAAA 89
    |||||||
Db 32 AACAAACCGACTCACAAAAAAAAAAAAAAAA 4
    |||||||

RESULT 6
CD364677/c
LOCUS
DEFINITION
  UI-H-FT2-bjnm-i-14-0-UI.s1 NCI CGAP FT2 Homo sapiens cDNA clone
  UI-H-FT2-bjnm-i-14-0-UI 3', mRNA sequence.
CD364677
CD364677.1 GI:31148767
EST.
  Homo sapiens (human)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Hominidae; Homo.
  1 (bases 1 to 691)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: csapsb@mail.nih.gov
  Tissue Procurement: Dr. Gary W. Hunninghake, U of I
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Distribution information can be found at
  http://genome.uiowa.edu/distribution/cgap.html
  Seq primer: M13 FORWARD
  POLYA=Yes.
FEATURES
  source      Location/Qualifiers
    1..691
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="UI-H-FT2-bjnm-i-14-0-UI"
      /tissue_type="Alveolar Macrophage"
      /dev_stage="Adult"

```

```

/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FT2"
/notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: Ecor I; Site 2: Not I;
NCI CGAP FT2 is a subtracted cDNA library constructed from
a pool of 81 RNA samples from Alveolar Macrophages
challenged with different treatments. The mRNA samples
were a mixture of these conditions (times refer to
incubations following isolation by bronchoalveolar lavage)
(some normal donor macrophages were cultured in some of
the conditions, other donor macrophages in different
conditions). The mRNA samples were pooled for library
construction. Control 0 hours; control 3 hours; control 24
hours; LPS 100 ng/ml, 3 hours; LPS 100 ng/ml, 24 hours;
PMA 10 ng/ml, 3 hours; PMA 10 ng/ml, 24 hours; Klebsiella
moi 10, 3 hours; Klebsiella moi 10, 24 hours; Staph
aureus moi 10, 3 hours; Staph aureus moi 10, 24 hours;
Adenoviral vector (Ad5 CMV eGFP), moi 500, 3 hours;
Adenoviral vector (Ad5 CMV eGFP), moi 500, 24 hours; wt
adenovirus moi 500, 3 hours; wt adenovirus moi 500, 24
hours; Ad vector + LPS 3 hours; Ad vector + LPS 24 hours;
wt adenovirus + LPS 3 hours; wt adenovirus + LPS 24 hours.
The library was subtracted according to Bonaldo, Lennon
and Soares, Genome Research, 6:791-806, 1996. The tissue
was provided by Dr. Gary W. Hunninghake of the University
of Iowa.
TAG_TISSUE=Human Lung Alveolar Macrophage
TAG_LIB=UI-H-FT2
TAG_SEQ=GGCCATGCCG"

ORIGIN
  Query Match      89.2%; Score 79.4; DB 6; Length 691;
  Best Local Similarity 93.3%; Pred. No. 8.5e-08;
  Matches 83; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGGGAATCCCCCTTCAGTCTTTTGAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
    |||||||
Db 91 CGGGAATCCCCCTTCAGTCTTTTGAAGTTCATGACTCGAATATCTGAAATGAAGAA 32
    |||||||

QY 61 AACAAACCAAAAAAAAAAAAAAAAAAGAAAA 89
    |||||||
Db 31 AACAAACCGACTCACAAAAAAAAAAAAAAAA 3
    |||||||

RESULT 7
BM053069
LOCUS
DEFINITION
  ie69h03.v3 Melton Normalized Human Islet 4 N4-His 1 Homo sapiens
  cDNA clone IMAGE:5672308 5', mRNA sequence.
BM053069
BM053069.1 GI:16809019
EST.
  Homo sapiens (human)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Hominidae; Homo.
  1 (bases 1 to 313)
  Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
  Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
  Hillier, L., Marra, M., Pape, D., Wyllie, T., Martin, J., Blistain, A.,
  Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
  Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarisvilli, R.,
  Williams, T., Jackson, Y. and Bowers, Y.
  Endocrine Pancreas Consortium
  Unpublished (2000)
  Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
  Endocrine Pancreas Consortium
  Harvard University, Howard Hughes Medical Institute
  Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
  MA 02138
  Tel: 617-495-1812
  Fax: 617-495-8557

```

Email: dmelton@biohp.harvard.edu
 Library was constructed by Dr. Douglas Melton DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Juliana Brown
 (brown@fas.harvard.edu) This sequence now available from the IMAGE
 consortium, for clone orders contact: info@image.llnl.gov.

FEATURES

source

1. 313
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5672308"
 /sex="Both"
 /tissue type="Islets of Langerhans"
 /dev stages="Adult"
 /lab host="DH10B"
 /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
 /note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;
 Site_2: Sal 1; Starting library constructed using
 SuperScript Plasmid Library kit (Life Technologies). cDNA
 made by oligo-dT priming. Size-selected by column
 fractionation; average insert size 1.08 kb. Library was
 amplified once on solid support and plasmid DNA from
 library was prepared. The library DNA was normalized by
 method #4 from Bonaldo, Lennon, and Soares 1996 Genome
 Research 6:791-806; 0.5 microgram single-stranded library
 Plasmid DNA was mixed with 5 micrograms PCR product
 representing library inserts and hybridized to an Ecot of
 20. Single-stranded (unhybridized) plasmids were isolated
 by hydroxyapatite chromatography and used to make this
 library."

ORIGIN

Query Match 84.7%; Score 75.4; DB 3; Length 313;
 Best Local Similarity 92.9%; Pred. No. 6.8e-07;
 Matches 79; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 Db 229 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 288
 |||||
 Qy 61 AACAAACCAAAAAAAGAA 85
 |||||
 Db 289 AACAAACCGACTCACAAAAAAA 313
 |||||

RESULT 8

BF431778/c
 LOCUS
 DEFINITION nabs0c04.x1 Soares NSF F8 9W OT PA_P_S1 mRNA linear EST 19-JAN-2001
 IMAGE:3269358 3', mRNA sequence.
 BF431778
 ACCESSION
 VERSION BF431778.1 GI:11443892
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40UP from Gibco

High quality sequence stop: 272.

Location/Qualifiers

1. 292

/organism="Homo sapiens"

FEATURES

source

/mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3269358"
 /lab host="DH10B"
 /clone_lib="Soares NSF F8 9W OT PA_P_S1"
 /note="Organ: pooled; Vector: pTT3D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
 Equal amounts of plasmid DNA from five normalized
 libraries were mixed, and ss circles were made in vitro.
 Following HAP purification, this DNA was used as tracer in
 a subtractive hybridization reaction. The driver was
 PCR-amplified cDNAs from pools of 5,000 clones made from
 the same 5 libraries. The pools consisted of the following
 libraries and cloneIDs: Soares NDHSF pool 1:
 309384-310919, 323208-325895 Soares NB2HP pool 1:
 145032-147335, 147720-148103, 148872-149255, 15002 -
 150407, 151176-152327 Soares NB2HP8-9W pool 1:
 758280-760583, 772104-774407 Soares NBHPA pool 1:
 304776-306311, 320136-322823, 326280-326663 Soares NBHOT
 pool 1: 723720-726407, 739080-740999 Subtraction by Bento
 Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 82.0%; Score 73; DB 2; Length 292;
 Best Local Similarity 93.8%; Pred. No. 2.3e-06;
 Matches 76; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
 |||||
 Db 81 CGGGAATCCCTTCAGTCTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 22
 |||||
 Qy 61 AACAAACCAAAAAAAGAA 81
 |||||
 Db 21 AACAAACCGACTCACAAAAAAA 1
 |||||

RESULT 9

AA903087/c

LOCUS

DEFINITION ok46d04.s1 NCI_CGAP_Lei2 Homo sapiens cDNA clone IMAGE:1516999 3',

mRNA sequence.

AA903087

VERSION AA903087.1 GI:3038210

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE 1 (bases 1 to 310)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

unknown library type

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 295.

Location/Qualifiers

1. 310

source

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:1516999"

/tissue type="leiomyosarcoma"

/lab host="DH10B"

/clone_lib="NCI CGAP Lei2"

/note="Organ: soft tissue; Vector: pTT3D-Pac (Pharmacia)

with a modified polylinker; Site_1: Not 1; Site_2: Eco RI;

1st strand cDNA was primed with a Not I - oligo(dT) primer

15'-AACTGGAAGATTCGCGCGCAATCGTTTTTTTTTTTTTT-3'],

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

```

ORIGIN
Query Match      82.0%; Score 73; DB 1; Length 310;
Best Local Similarity 88.8%; Pred. No. 2.3e-06;
Matches 79; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CGGAATCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTCGAATGAAGAA 60
Db 90 CGGAATCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTCGAATGAAGAA 31
Qy 61 AACAAACCAAAAAAAGAAAAAAGAAAAA 89
Db 30 AACAAACCGACTCAAAACCTCAAAAAA 2

RESULT 10
CK903979/c
LOCUS
DEFINITION
i69h03.x5 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA clone IMAGE:5672308 3', mRNA sequence.
ACCESSION
CK903979
VERSION
CK903979.1 GI:45365510
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 339)
Melton,D., Meadows,A., Clifton,S., Hillier,L., Marra,M., Pape,D.,
Wylie,T., Martin,J., Blisstein,A., Schmitt,A., Theising,B.,
Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M.,
McCann,R., Cole,R., Tagareishvili,R., Williams,T., Jackson,Y. and
Bowers,Y.
WashU-Harvard Pancreas EST Project
Unpublished (2000)
Other ESTs: ie69h03.y3
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
This read is a 3' RESEQUENCE of a previously sequenced pancreas
clone
This resequenced clone has not previously been sequenced on this
end, resequencing from this end represents new data
Seq primer: -40UP from Gibco.

FEATURES
source
1..339
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5672308"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/note="Organ: Pancreas; Vector: pSPORT1; Site 1: Not 1;
Site 2: Sal 1; Starting library constructed using
SuperScript plasmid library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome

```

Research 6:791-806; 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an EcoI of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

```

ORIGIN
Query Match      82.0%; Score 73; DB 7; Length 339;
Best Local Similarity 93.8%; Pred. No. 2.3e-06;
Matches 76; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAATCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTCGAATGAAGAA 60
Db 81 CGGAATCCCTTCAGTCTTTGAAAAAGTTCATGACTCGAATATCTCGAATGAAGAA 22
Qy 61 AACAAACCAAAAAAAGAAAAAAGAAAAA 81
Db 21 AACAAACCGACTCAAAAAA 1

RESULT 11
AI927624/c
LOCUS
DEFINITION
w085a02.x1 NCI_CGAP_kid11 Homo sapiens cDNA clone IMAGE:2462090 3',
mRNA sequence.
ACCESSION
AI927624
VERSION
AI927624.1 GI:5663588
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 475)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb-rc@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 1810 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 456.

FEATURES
source
1..475
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2462090"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid11"
/note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP Kid3 was
prepared, and ss circles were made in vitro. Following HAP
hybridization, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."

ORIGIN
Query Match      82.0%; Score 73; DB 1; Length 475;

```

Best Local Similarity 93.8%; Pred. No. 2.2e-06;
Matches 76; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
|||||
Db 81 CGGNATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 22
|||||
Qy 61 AACAAACCAAAAAA 81
|||||
Db 21 AACAAACCGACTCACAAAAA 1
|||||

RESULT 12
BF002331/c
LOCUS
DEFINITION 716 bp mRNA linear EST 06-OCT-2000
7h22b08.x1 NCI CGAP Col6 Homo sapiens cDNA clone IMAGE:3316695 3',
similar to TR:O00466 O00466 K12 PROTEIN PRECURSOR. ;, mRNA
sequence.

ACCESSION BF002331 GI:10702606
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 716)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov

Seq primer: -40UP from Gibco
High quality sequence stop: 462.

FEATURES

source
Location/Qualifiers
1..716
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3316695"
/tissue_type="colon tumor, RER"
/lab_host="DH10B"
/clone_lib="NCI CGAP Col6"

/note="Organ: colon; Vector: p7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Col6 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1057416-1061255, and 1144584-1145351).
Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 80.9%; Score 72; DB 2; Length 716;
Best Local Similarity 93.8%; Pred. No. 3.6e-06;
Matches 75; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
|||||
Db 80 CGGNATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 21
|||||
Qy 61 AACAAACCAAAAAA 80
|||||

Db 20 AACAAACCGACTCACAAAAA 1
|||||

RESULT 13
AI480347/c
LOCUS
DEFINITION 445 bp mRNA linear EST 14-APR-1999

AI480347 tms1c09.x1 NCI_CGAP Kid11 Homo sapiens cDNA clone IMAGE:2161648 3',
mRNA sequence.

ACCESSION AI480347
VERSION AI480347.1 GI:4373515
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 445)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 533 Std Error: 0.00
Seq primer: -40UP from Gibco.

Location/Qualifiers
1..445

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2161648"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid11"

/note="Organ: kidney; Vector: p7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."

FEATURES
source
Location/Qualifiers
1..445

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2161648"
/lab_host="DH10B"

/note="Organ: kidney; Vector: p7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."

ORIGIN
Query Match 80.4%; Score 71.6; DB 1; Length 445;
Best Local Similarity 94.9%; Pred. No. 4.6e-06;
Matches 74; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 60
|||||

Db 78 CGGGAATCCCTTCAGTCTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGAAGAA 19
|||||

Qy 61 AACAAACCAAAAAA 78
|||||

Db 18 AACAAACCAAAAAA 1
|||||

RESULT 14
BU689727/c
LOCUS
DEFINITION 687 bp mRNA linear EST 07-OCT-2002

UI-CF-FNO-aet-k-09-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone
UI-CF-FNO-aet-k-09-0-UI 3', mRNA sequence.

ACCESSION BU689727
VERSION BU689727.1 GI:23547747

KEYWORDS EST. Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
AUTHORS 1 (bases 1 to 687)
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
PUBMED 8889548
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 286-351, >GC rich#Low_complexity
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES source
Location/Qualifiers
1..687
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="UI-CF-FNO-aet-k-09-0-UI"
/tissue_types="Human Lung Epithelial cells"
/lab_host="DH10B (Life Technologies) (TI phage resistant)"
/clone_lib="UI-CF-FNO"
/notes="Organ: Lung; Vector: p7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-FNO is a subtracted cDNA library derived from two
normalized Human lung epithelial cell libraries (EN1 and
DUI). The library was subtracted according to according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. For additional information, contact:
bento-soares@uiowa.edu
TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS
6hr to LPS 24h
TAG LIB=UI-CF-FNO
TAG_SEQ=CTGCTCAGGT"

ORIGIN
Query Match 80.2%; Score 71.4; DB 5; Length 687;
Best Local Similarity 87.6%; Pred. No. 4.9e-06;
Matches 78; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy 1 CGGGAATCCCCCTTCAGTCTTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
Db 101 CGGGAATCCCCCTTCAGTCTTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 42
Qy 61 AACAAACCAAAAAAAAAAAAAAAAAA 89
Db 41 AACAAACCGACTCAAAACCTCCAATAA 13

RESULT 15
AA826896/c
LOCUS
DEFINITION nr86c10.s1 NCI_CGAP_Pr24 Homo sapiens cDNA clone IMAGE:1174866 3',
mRNA sequence.
AA826896
ACCESSION
VERSION AA826896.1 GI:2900893

KEYWORDS EST. Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 109)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Suzanne L. Topalian, M.D., Robert K. Bright,
Ph.D.
CDNA Library Preparation: Stratagene, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 1235 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 104.
Location/Qualifiers
1..109
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1174866"
/tissue_types="invasive tumor (cell line)"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI_CGAP_Pr24"
/notes="Organ: prostate; Vector: Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Invasive prostate tumor cell line (HPV
immortalized). 5' adaptor sequence: 5' GAATTCGACGAG 3'
3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'
Average insert size: 1.0 kb."

ORIGIN
Query Match 78.7%; Score 70; DB 1; Length 109;
Best Local Similarity 93.6%; Pred. No. 1.1e-05;
Matches 73; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CGGGAATCCCCCTTCAGTCTTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 60
Db 82 CGGGAATCCCCCTTCAGTCTTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAA 23
Qy 61 AACAAACCAAAAAAAAAA 78
Db 22 AACAAACCGACTCACAAA 5

Search completed: December 7, 2005, 09:55:48
Job time : 1015.97 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:34:11 ; Search time 60.7533 Seconds
(without alignments)
455.919 Million cell updates/sec

Title: US-09-980-046B-4
Perfect score: 89
Sequence: 1 cgggaatcccccttcagtc.....aaaaaaaaaagaaaaa 89

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New: *
1: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq: *
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq: *
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq: *
4: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq: *
5: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq: *
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq: *
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq: *
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq: *
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq: *
10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41.4	46.5	2066	US-10-523-588-2	Sequence 2, Appli
2	36.8	41.3	2027	US-10-131-826A-175	Sequence 175, App
3	36.2	40.7	2668	US-10-131-826A-511	Sequence 511, App
4	36	40.4	1941	US-10-131-826A-165	Sequence 165, App
5	36	40.4	2625	US-11-102-978-12	Sequence 12, Appl
6	35.8	40.2	1872	US-10-750-185-31235	Sequence 31235, A
7	35.2	39.6	171427	US-11-112-908-60	Sequence 60, Appl
8	35	39.3	644	US-11-102-240-55	Sequence 55, Appl
9	34.6	38.9	778	US-10-986-501-38	Sequence 38, Appl
10	34.6	38.9	5503	US-11-112-908-6	Sequence 6, Appli
11	34.4	38.7	1968	US-10-131-826A-163	Sequence 163, App
12	34.4	38.7	2537	US-11-102-978-8	Sequence 8, Appli
13	34.4	38.7	3124	US-10-750-185-38847	Sequence 38847, A
14	34.2	38.4	1478	US-10-909-125-1744	Sequence 1744, Ap
15	34.2	38.4	4740	US-11-080-026-3	Sequence 3, Appli
16	34	38.2	485	US-11-102-240-43	Sequence 43, Appl
17	33.8	38.0	687	US-10-986-501-107	Sequence 107, App
18	33.8	38.0	2091	US-10-276-233A-19	Sequence 19, Appl
19	33.8	38.0	2933	US-10-131-826A-345	Sequence 345, App
20	33.8	38.0	3501	US-10-131-826A-37	Sequence 37, Appl
21	33.6	37.8	783	US-10-131-826A-321	Sequence 321, App
22	33.6	37.8	1979	US-10-967-527A-31	Sequence 31, Appl
23	33.4	37.5	2037	US-10-909-125-799	Sequence 799, App

ALIGNMENTS

RESULT 1

US-10-523-588-2
; Sequence 2, Application US/10523588
; Publication No. US20050251870A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: CSNK1GS AS MODIFIERS OF THE P21 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX03-057C-US
; CURRENT APPLICATION NUMBER: US/10/523,588
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: US 60/401,739
; PRIOR FILING DATE: 2002-08-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 2066
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-523-588-2

Query Match 46.5%; Score 41.4; DB 6; Length 2066;
Best Local Similarity 68.7%; Pred. No. 0.13;
Matches 57; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy	7	TCCCCCTTCAGTCTCTTTGAAAAAGCTTCATGACTCGAATATCTGAATGTGAAGAAACAAA	66
Db	1953	TGCACCTTTTGTAGTTTTTTTAAAAAAAACAAAACAAACATGGCGATGTCACAAAAA	2012
Qy	67	CCAAAAAAGAAAAAGAAAAA	89
Db	2013	AAAAAAGAAAAAAGAAAAA	2035

RESULT 2

US-10-131-826A-175
; Sequence 175, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Deanoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey


```
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131.826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 165
; LENGTH: 1941
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-131-826A-165

Query Match 40.4%; Score 36; DB 6; Length 1941;
Best Local Similarity 67.1%; Pred. No. 2.1;
Matches 51; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 14 TCAGTCTCTTGAAGAAAGTTCCTGACTCGAATATCTGAATCTGAATCAAGAAACCAACCAAAAA 73
Db 1860 TCTGTGACTGATTAAGTTCATTTTGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1919

Qy 74 AAAAAAGAAAGAAAGAAAG 89
Db 1920 AAAAAAGAAAGAAAGAAAG 1935

RESULT 5
US-11-102-978-12
; Sequence 12, Application US/11102978
; Publication No. US20050250142A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Technology Transfer Office
; APPLICANT: University of Utah Research Foundation
; TITLE OF INVENTION: Diagnosis and Treatment of Herpes Simplex Virus Disease
; FILE REFERENCE: 0274-5537.1US
; CURRENT APPLICATION NUMBER: US/11/102,978
; CURRENT FILING DATE: 2005-04-11
; PRIOR APPLICATION NUMBER: PCT/US2003/033152
; PRIOR FILING DATE: 2003-10-18
; PRIOR APPLICATION NUMBER: 60/419,576
; PRIOR FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 2625
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(2625)
; OTHER INFORMATION: CHODL
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (392)..(1213)
```

```
; OTHER INFORMATION: The gene is CHODL, encoding chondrolectin, NM024944.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (506)..(928)
; OTHER INFORMATION: C-type lectin (CTL) or carbohydrate-recognition domain (CRD)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (539)..(931)
; OTHER INFORMATION: C-type lectin domain (CRD), including both the long and short
; OTHER INFORMATION: form C-type.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1975)..(1975)
; OTHER INFORMATION: Comprises A or G.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2032)..(2032)
; OTHER INFORMATION: Comprises A or G.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2121)..(2121)
; OTHER INFORMATION: Comprises A or T.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2349)..(2349)
; OTHER INFORMATION: Comprises G or T.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2403)..(2404)
; OTHER INFORMATION: TCTA
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2543)..(2543)
; OTHER INFORMATION: Comprises A or G.
; US-11-102-978-12

Query Match 40.4%; Score 36; DB 7; Length 2625;
Best Local Similarity 64.3%; Pred. No. 2.1;
Matches 54; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 6 ATCCCTTCAGTCTCTTGAAGAAAGTTCCTGACTCGAATATCTGAATCAAGAAAGAAAGAAAGAAAG 65
Db 2496 ATGTCCTCTGCTCTCTTTTAAACCAATAAAGAGTCTTGTGTTCTGAAGAAAGAAAGAAAGAAAG 2555

Qy 66 ACCAAAGAAAGAAAGAAAGAAAGAAAG 89
Db 2556 AAAAAAGAAAGAAAGAAAGAAAGAAAG 2579

RESULT 6
US-10-750-185-31235
; Sequence 31235, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: Denise, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31235
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Bovine 19866881231610
```

US-10-750-185-31235

Query Match 40.2%; Score 35.8; DB 6; Length 1872;
Best Local Similarity 63.2%; Pred. No. 2.3;
Matches 55; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
Qy 3 GGAAATCCCTTCAGTTCTTTGAAAGAGTTCATGACTCGAATATCTGAAATGAAGAAAA 62
Db 80 GAAAGACAGTAATAATTTTAAAGAAAGATCAACCAAGAAATTAACAATTTGAAACACA 139
Qy 63 CAAACCAAAAAAAGAAAAA 89
Db 140 AAGAGAAAAAAGAGAAAAA 166

RESULT 7

US-11-112-908-60/c
; Sequence 60, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-23
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 60
; LENGTH: 171427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-60

Query Match 39.6%; Score 35.2; DB 7; Length 171427;
Best Local Similarity 76.8%; Pred. No. 4.5;
Matches 43; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Qy 34 CATGACTCGAATATCTCGAATGAAGAAACCAACCAAAAAAAGAAAAA 89
Db 104488 CAAGAATCCCATTTTCAGAAAAAAGAAAAAAGAAAAAAGAAAAA 104433

RESULT 8

US-11-102-240-55
; Sequence 55, Application US/11102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS
; FILE REFERENCE: P3230R1C106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: 10/063662
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262

; PRIOR FILING DATE: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 55
; LENGTH: 644
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-102-240-55

Query Match 39.3%; Score 35; DB 7; Length 644;
Best Local Similarity 63.9%; Pred. No. 3.2;
Matches 53; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Qy 7 TCCCTTCAGTTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAAAAA 66
Db 532 TCCACTGAATTCCTCATGAATCAAACTTATTGACACAGCAAAAAA 591

Qy 67 CCAAAAAAAGAAAAA 89
Db 592 AAAAAAAGAAAAA 614

RESULT 9

US-10-986-501-38
; Sequence 38, Application US/10986501
; Publication No. US20050244845A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 90 Human Secreted Proteins
; FILE REFERENCE: P2013P2C1
; CURRENT APPLICATION NUMBER: US/10/986,501
; CURRENT FILING DATE: 2004-11-12
; PRIOR APPLICATION NUMBER: US/10/621,363
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: 09/969,730
; PRIOR FILING DATE: 2001-10-06
; PRIOR APPLICATION NUMBER: 09/774,639
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 60/238,291
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/244,112
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: PCT/US98/16235
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/056,371
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,732
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,366
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 60/056,364
; PRIOR FILING DATE: 1997-08-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 373
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 778
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-986-501-38

Query Match 38.9%; Score 34.6; DB 6; Length 778;
Best Local Similarity 67.1%; Pred. No. 3.9;
Matches 49; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 17 GTTCTTTGAAAAAGTTCATGACTCGAATATCTGAAATGAAGAAAAA 76
Db 695 GTCATTGTAATAAAGCCTCTGCTCTCAGATGTAAAAAAGAAAAA 754
Qy 77 AAAAAAAGAAAA 89
Db 755 AAAAAAAGAAAA 767

```

; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 163
; LENGTH: 1968
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-131-826A-163

Query Match      38.7%; Score 34.4; DB 6; Length 1968;
Best Local Similarity 63.1%; Pred. No. 4.7;
Matches 53; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY      6  ATCCCCCTTCAGTCTTTTGAAAAAGTTCCATGACTCGAATATCTGAAATGGAAGAAAAACAA 65
Db      1883  ATCTATTTCATCTCGAAAAATAAGAATATATATTTCACTCTAAAAA
QY      66  ACCAAAAA
Db      1943  AAAAAA
; US-11-102-978-8
RESULT 12
US-11-102-978-8
; Sequence 8, Application US/11102978
; Publication No. US20050250142A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Technology Transfer Office
; APPLICANT: University of Utah Research Foundation
; TITLE OF INVENTION: Diagnosis and Treatment of Herpes Simplex Virus Disease
; FILE REFERENCE: 0274-5537.1US
; CURRENT APPLICATION NUMBER: US/11/102,978
; CURRENT FILING DATE: 2005-04-11
; PRIOR APPLICATION NUMBER: PCT/US2003/033152
; PRIOR FILING DATE: 2003-10-18
; PRIOR APPLICATION NUMBER: 60/419,576
; PRIOR FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 2537
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (102)..(1199)
; OTHER INFORMATION: Homo sapiens coxsackie virus and adenovirus receptor (CXADR)
; OTHER INFORMATION: mRNA, genbank NM001338.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (186)..(515)
; OTHER INFORMATION: IG, Immunoglobulin
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (207)..(467)
; OTHER INFORMATION: Igv, Immunoglobulin V-Type
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (558)..(758)
; OTHER INFORMATION: Igc2, Immunoglobulin C-2 Type

```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (564)..(788)
; OTHER INFORMATION: IG, Immunoglobulin
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (750)..(750)
; OTHER INFORMATION: includes an A or G nucleotide
; FEATURE:
; NAME/KEY: allele
; LOCATION: (969)..(969)
; OTHER INFORMATION: may comprise G or A.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1148)..(1148)
; OTHER INFORMATION: may comprise G or A.
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (1148)..(1148)
; OTHER INFORMATION: Comprises A or C.
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (1210)..(1210)
; OTHER INFORMATION: Comprises T or C.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1269)..(1269)
; OTHER INFORMATION: may comprise G or T.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1278)..(1278)
; OTHER INFORMATION: may comprise G or T.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1293)..(1293)
; OTHER INFORMATION: may comprise G or A.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1297)..(1297)
; OTHER INFORMATION: may comprise G or A.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1298)..(1298)
; OTHER INFORMATION: may comprise C or T.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1322)..(1322)
; OTHER INFORMATION: may comprise A or G.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1365)..(1365)
; OTHER INFORMATION: may comprise A or G.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1376)..(1376)
; OTHER INFORMATION: may comprise A or G.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1381)..(1381)
; OTHER INFORMATION: may comprise C or T.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1397)..(1397)
; OTHER INFORMATION: may comprise C or A.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1449)..(1449)
; OTHER INFORMATION: may comprise G or T.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (1827)..(1827)
; OTHER INFORMATION: may comprise G or A.
; FEATURE:

; NAME/KEY: allele
; LOCATION: (1866)..(1866)
; OTHER INFORMATION: may comprise C or T.
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (1976)..(1976)
; OTHER INFORMATION: Comprises A or G.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2007)..(2007)
; OTHER INFORMATION: may comprise C or T.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2095)..(2095)
; OTHER INFORMATION: may comprise A or G.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2106)..(2106)
; OTHER INFORMATION: may comprise A or C.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2154)..(2154)
; OTHER INFORMATION: may comprise T or C.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2159)..(2159)
; OTHER INFORMATION: may comprise T or C.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2192)..(2192)
; OTHER INFORMATION: may comprise T or C.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2193)..(2193)
; OTHER INFORMATION: may comprise A or G.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2271)..(2271)
; OTHER INFORMATION: may comprise A or C.
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (2294)..(2294)
; OTHER INFORMATION: Comprises T or C.
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (2328)..(2328)
; OTHER INFORMATION: Comprises T or C.
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (2340)..(2340)
; OTHER INFORMATION: Comprises T or A.
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (2346)..(2346)
; OTHER INFORMATION: Comprises G or A.
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (2354)..(2354)
; OTHER INFORMATION: Comprises G or A.
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (2374)..(2374)
; OTHER INFORMATION: Comprises G or A.
; FEATURE:
; NAME/KEY: Unsure
; LOCATION: (2413)..(2413)
; OTHER INFORMATION: Comprises C or T.
; FEATURE:
; NAME/KEY: allele
; LOCATION: (2440)..(2440)
; OTHER INFORMATION: may comprise A or G.
; FEATURE:

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 778.261 Seconds
(without alignment)
6427.437 Million cell updates/sec

Title: US-09-980-046B-5
Perfect score: 88
Sequence: 1 cggagtcgagtcctgtatca.....cgtgttcttgggtacaaaaa 88

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_in.*
- 3: gb_env.*
- 4: gb_ov.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pt.*
- 9: gb_ro.*
- 10: gb_scs.*
- 11: gb_by.*
- 12: gb_un.*
- 13: gb_vl.*
- 14: gb_brg.*
- 15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81.6	92.7	645	8	BC022439
2	80	90.9	251	10	BV167599
3	80	90.9	467	6	CQ415312
4	80	90.9	536	6	CQ433071
5	80	90.9	541	6	CQ425463
6	80	90.9	630	8	BC070243
7	80	90.9	642	6	BD237072
8	80	90.9	642	6	AR225472
9	80	90.9	642	6	AR562895
10	80	90.9	642	6	AX321588
11	80	90.9	222085	14	AC144988
12	78.4	89.1	637	8	BC006794
13	78.4	89.1	695	6	AX281855
14	78.4	89.1	749	8	BC008417
15	76.8	87.3	451	6	BD222194
16	76.8	87.3	451	6	AX014315
17	75.8	86.1	118	6	CQ670675
18	75.8	86.1	594	6	CQ923967

c	19	75.8	86.1	668	6	CQ434132	CQ434132 Sequence
	20	75.8	86.1	808	6	BD190297	BD190297 Elkl phos
	21	75.8	86.1	808	6	CQ812348	CQ812348 Sequence
	22	75.8	86.1	808	6	CS036376	CS036376 Sequence
	23	75.8	86.1	808	6	CS045328	CS045328 Sequence
	24	75.8	86.1	808	6	CS073286	CS073286 Sequence
	25	75.8	86.1	808	6	CS108221	CS108221 Sequence
	26	75.8	86.1	808	6	AR380734	AR380734 Sequence
	27	75.8	86.1	808	6	AR489210	AR489210 Sequence
	28	75.8	86.1	808	6	HS18U	X57352 Human 1-8U
	29	75.8	86.1	143779	8	AC136475	AC136475 Homo sapi
	30	75.8	86.1	156819	14	AP006283	AP006283 Homo sapi
	31	75.8	86.1	175416	14	AP006286	AP006286 Homo sapi
	32	75.8	86.1	187160	8	AC138230	AC138230 Homo sapi
	33	75.4	85.7	401	6	CQ695754	CQ695754 Sequence
	34	75.4	85.7	606	6	BD190296	BD190296 Elkl phos
	35	75.4	85.7	621	6	BD190298	BD190298 Elkl phos
	36	74.4	84.5	1933	6	E01979	E01979 DNA encoding
	37	73.6	83.6	422	10	AB146371	AB146371 Homo sapi
	38	73.6	83.6	84979	8	AC116049	AC116049 Homo sapi
	39	73.6	83.6	176104	14	AC016193	AC016193 Homo sapi
	40	73.6	83.6	185237	8	AC011124	AC011124 Homo sapi
	41	73.6	83.6	189572	14	AC084812	AC084812 Homo sapi
	42	72.6	82.5	47165	8	HSDJ781L3	AL121994 Human DNA
	43	72.6	82.5	204096	14	AC027704	AC027704 Homo sapi
	44	72.4	82.3	346	6	CQ661861	CQ661861 Sequence
	45	72	81.8	51649	8	BX927182	BX927182 Human DNA

ALIGNMENTS

BC022439 645 bp mRNA linear PRI 29-JUN-2004
Homo sapiens interferon induced transmembrane protein 3 (1-80),
mRNA (cdna clone MGC:24755 IMAGE:4282809), complete cds.

BC022439

BC022439.1 GI:18490258

MGC.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 645)

Klausner,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Shat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Udén,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
McKernan,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettner,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilov,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 645)

Strausberg,R.

Direct Submission

Submitted (01-FEB-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

REMARK
COMMENT
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapsb-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 35 Row: m Column: 16.

FEATURES

source
 1. .645
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:24755 IMAGE:4282809"
 /tissue type="Brain, primitive neuroectodermal"
 /clone_lib="NIH_MGC_56"
 /lab_host="DH10B"
 /note="vector: pDNR-LIB"
 1. .645
 /gene="IFITM3"
 /note="synonym: 1-8U"
 /db_xref="GeneID:10410"
 /db_xref="MIM:605579"
 51. .452
 /gene="IFITM3"
 /codon_start=1
 /product="interferon-induced transmembrane protein 3 (1-8U)"
 /protein_id="AAH22439.1"
 /db_xref="GI:18490259"
 /db_xref="GeneID:10410"
 /db_xref="MIM:605579"
 /translation="MNQTVTFSPVNSGPPNYEMLKEHEVAVLGAPHPAPPTST
 VIHRTSPVDHVWLSFNTLWNPCLGFIAPYVKSRRDKWGDVTGAQAYAST
 AKCLNIWALILGILMTLLIVPLVIFQYQ"

gene

CDS

ORIGIN

Query Match 92.7%; Score 81.6; DB 8; Length 645;
 Best Local Similarity 95.5%; Pred. No. 6.6e-17;
 Matches 84; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCACAGCTTTTCTACAATGGCATTCAAT 60
 Db 535 CGGAGCGGAGTCCTGTATCAGCCCTTTATCTCACAGCTTTTCTACAATGGCATTCAAT 594
 Qy 61 AAAGTGCACGTGTTCTTGGTACAAAAA 88
 Db 595 AAAGTGCACGTGTTCTTGGTAAAAAAA 622
 RESULT 2
 BV167599/c
 LOCUS sqm5991 Human DNA (Sequenom) Homo sapiens STS 10-JUN-2004
 DEFINITION tagged site.
 ACCESSION BV167599
 VERSION BV167599.1 GI:48001170
 KEYWORDS STS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE

1 (bases 1 to 251)
 Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
 Cantor,C.R. and Braun,A.
 Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
 Regions
 Genome Res. (2004) In press

COMMENT

Contact: Andreas Braun
 Pharmaceuticals division
 Sequenom, Inc.
 3595 John Hopkins Court, San Diego, CA 92121, USA
 Tel: 18582029018
 Fax: 18582029020
 Email: abraun@sequenom.com
 Primer A: No primer sequence submitted
 Primer B: No primer sequence submitted
 STS size: 251.

FEATURES

source
 1. .251
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone_lib="Human DNA (Sequenom)"
 <1. .>251

ORIGIN

Query Match 90.9%; Score 80; DB 10; Length 251;
 Best Local Similarity 94.3%; Pred. No. 2.5e-16;
 Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCACAGCTTTTCTACAATGGCATTCAAT 60
 Db 107 CGGAGCGGAGTCCTGTATCAGCCCTTTATCTCACAGCTTTTCTACAATGGCATTCAAT 48
 Qy 61 AAAGTGCACGTGTTCTTGGTACAAAAA 88
 Db 47 AAAGTGCACGTGTTCTTGGWAAAAAAA 20
 RESULT 3
 CQ415312
 LOCUS
 DEFINITION Sequence 346 from Patent WO0151628.
 ACCESSION CQ415312
 VERSION CQ415312.1 GI:41367541
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE

1
 Lillie,J., Xu,Y., Wang,Y. and Steinmann,K.
 Novel genes, compositions, kits, and methods for identification,
 assessment, prevention, and therapy of breast cancer
 Patent: WO 0151628-A 346 19-JUL-2001;
 Millennium Pharmaceuticals, Inc. (US)

FEATURES

source
 1. .467
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 90.9%; Score 80; DB 6; Length 467;
 Best Local Similarity 94.3%; Pred. No. 2.4e-16;
 Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCACAGCTTTTCTACAATGGCATTCAAT 60
 Db 67 CGGAGCGGAGTCCTGTATCAGCCCTTTATCTCACAGCTTTTCTACAATGGCATTCAAT 126
 Qy 61 AAAGTGCACGTGTTCTTGGTACAAAAA 88


```

1  CCGAGTTCGAGTCCGTATCAGCCCTTTATCTCACACGCTTTTCTACAAATGCGATTCAAAT  60
/mol_type="mrna"
/db_xref="taxon:9606"

```

```
/c1one="MGC:88228 IMAGE:30401041"
/tissue_type="Blood vessels- blood vessel, aorta, basilar
artery."
/clone_lib="NIH_MGC_187"
/lab_host="DH10B"
/notes="Vector: pDNR-LIB"
1. .630
/gene="IFITM3"
/notes="synonym: 1-8U"
/db_xref="GeneID:10410"
/db_xref="MIM:605579"
36. .437
/gene="IFITM3"
/codon_start=1
/product="interferon-induced transmembrane protein 3
(1-8U)"
/protein_id="AAH70243.1"
/db_xref="GI:47682387"
/db_xref="GeneID:10410"
/db_xref="MIM:605579"
/translation="MNTVQTFSPVNSGPPNYEMLKESHEVAVLGAHPNPPTST
VIHRSVDPDHVWSLENTLFNPPCLGFIAFAYSVKSRDRRMVDVGTGAQAYAST
AKCLNIWALILGILMTILLIPIVLPFAQY"
```

ORIGIN

```
Query Match 90.9%; Score 80; DB 8; Length 630;
Best Local Similarity 94.3%; Pred. No. 2.3e-16;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCCTCACAGCTTTTCTACAATGGCATTCAAT 60
|||||
Db 520 CGGAGCGAGTCCTGTATCAGCCCTTTATCCTCCTCACAGCTTTTCTACAATGGCATTCAAT 579
|||||

Qy 61 AAAGTGCACGTGTTCTTCTGGTGACAAAAA 88
|||||
Db 580 AAAGTGCACGTGTTCTTCTGATATAAAAAA 607
|||||
```

RESULT 7

```
BD237072
LOCUS BD237072 Compounds for remedy and diagnosis of lung cancer and method for
DEFINITION using the same. 642 bp DNA linear PAT 17-JUL-2003
ACCESSION BD237072.1 GI:33046842
VERSION JP 2002516659-A/73.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 642)
REFERENCE Reed,S.G., Lodes,M.J., Frudakis,T.N. and Mohamath,R.
AUTHORS Compounds for remedy and diagnosis of lung cancer and method for
TITLE using the same
JOURNAL Patent: JP 2002516659-A 73 11-JUN-2002;
CORIXA CORP
```

COMMENT

```
OS Homo sapiens (human)
PN JP 2002516659-A/73
PD 11-JUN-2002
PF 26-JAN-1999 JP 2000529432
PR 28-JAN-1998 US 09/015029,28-JAN-1998 US 09/015022 PR
18-MAR-1998 US 09/040828,18-MAR-1998 US 09/040831 PR
23-JUL-1998 US 09/122192,23-JUL-1998 US 09/122191 PR
PI STEVEN G REED,MICHAEL J LODES,TONY N FRUDAKIS,RAODOH MOHAMATH
PC C12N15/09,A61K35/14,A61K38/00,A61K39/00,A61K39/39,A61K39/395,
PC A61K39/395,
PC A61F11/00,A61P35/00,C07K14/47,C07K16/18,C07K19/00,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12P21/08,C12Q1/68,G01N33/53,G01N33/574,G01N33/577//
PC (C12N1/21,C12R1:19),C12N15/00,A61K37/02,C12N5/00 CC
Compounds for remedy and diagnosis of lung cancer and method CC
```

```
CC same for using the
FH Key Location/Qualifiers
FT source 1. .642
/organism="Homo sapiens (human)"
FEATURES
source 1. .642
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
```

ORIGIN

```
Query Match 90.9%; Score 80; DB 6; Length 642;
Best Local Similarity 94.3%; Pred. No. 2.3e-16;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCCTCACAGCTTTTCTACAATGGCATTCAAT 60
|||||
Db 535 CGGAGCGAGTCCTGTATCAGCCCTTTATCCTCCTCACAGCTTTTCTACAATGGCATTCAAT 594
|||||

Qy 61 AAAGTGCACGTGTTCTTCTGGTGACAAAAA 88
|||||
Db 595 AAAGTGCACGTGTTCTTCTGGTGACAAAAA 622
|||||
```

RESULT 8

```
AR225472
LOCUS AR225472 642 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 119 from patent US 6444425.
ACCESSION AR225472
VERSION AR225472.1 GI:27263418
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 642)
AUTHORS Reed,S.G., Lodes,M.J., Mohamath,R. and Secrist,H.
TITLE Compounds for therapy and diagnosis of lung cancer and methods for
their use
JOURNAL Patent: US 6444425-A 119 03-SEP-2002;
CORIXA Corporation; Seattle, WA;
WOX;
```

```
FEATURES
source 1. .642
Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"
```

ORIGIN

```
Query Match 90.9%; Score 80; DB 6; Length 642;
Best Local Similarity 94.3%; Pred. No. 2.3e-16;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCCTCACAGCTTTTCTACAATGGCATTCAAT 60
|||||
Db 535 CGGAGCGAGTCCTGTATCAGCCCTTTATCCTCCTCACAGCTTTTCTACAATGGCATTCAAT 594
|||||

Qy 61 AAAGTGCACGTGTTCTTCTGGTGACAAAAA 88
|||||
Db 595 AAAGTGCACGTGTTCTTCTGGTGACAAAAA 622
|||||
```

RESULT 9

```
AR562895
LOCUS AR562895 642 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 119 from patent US 6759508.
ACCESSION AR562895
VERSION AR562895.1 GI:53977724
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 642)
AUTHORS Lodes,M.J., Mohamath,R., Henderson,R.A., Benson,D.R. and Secrist,H.
```

```

TITLE      Compositions and methods for the therapy and diagnosis of lung
JOURNAL    Cancer
FEATURES   Patent: US 679508-A 119 06-JUL-2004;
            Corixa Corporation; Seattle, WA
            Location/Qualifiers
            1..642
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match      90.9%; Score 80; DB 6; Length 642;
Best Local Similarity 94.3%; Pred. No. 2.3e-16;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGATCAGCCCTTTATCTCCTCACAGCCTTTCTACAAATGGCATTCAAT 60
    |||||
Db 535 CGGAGCCGAGTCCTGATCAGCCCTTTATCTCCTCACAGCCTTTCTACAAATGGCATTCAAT 594
    |||||

Qy 61 AAAGTGCACGTGTTCTTCTGGTACAAAA 88
    |||||
Db 595 AAAGTGCACGTGTTCTTCTGGTAAAAAA 622

RESULT 10
AX321588
LOCUS      AX321588                642 bp      DNA      linear      PAT 15-DEC-2001
DEFINITION Sequence 119 from Patent WO0172295.
ACCESSION  AX321588
VERSION    AX321588.1  GI:17905704
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM  Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.

REFERENCE
1
AUTHORS   Reed,S.G., Lodes,M.J., Mohamath,R., Secrist,H., Benson,D.R.,
            Indrias,C.Y., Henderson,R.A., Fling,S.P., Algate,P.A., Elliot,M.,
            Mannion,J. and Kalos,M.D.
TITLE     Compositions and methods for the therapy and diagnosis of lung
JOURNAL  Cancer
FEATURES  Patent: WO 0172295-A 119 04-OCT-2001;
            CORIXA CORPORATION (US)
            Location/Qualifiers
            1..642
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match      90.9%; Score 80; DB 6; Length 642;
Best Local Similarity 94.3%; Pred. No. 2.3e-16;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGATCAGCCCTTTATCTCCTCACAGCCTTTCTACAAATGGCATTCAAT 60
    |||||
Db 535 CGGAGCCGAGTCCTGATCAGCCCTTTATCTCCTCACAGCCTTTCTACAAATGGCATTCAAT 594
    |||||

Qy 61 AAAGTGCACGTGTTCTTCTGGTACAAAA 88
    |||||
Db 595 AAAGTGCACGTGTTCTTCTGGTAAAAAA 622

RESULT 11
AC144988/c
LOCUS      AC144988                222085 bp      DNA      linear      HTG 10-JUL-2003
DEFINITION Gorilla gorilla clone CH255-209M14, WORKING DRAFT SEQUENCE,
            8 ordered pieces.
ACCESSION  AC144988
VERSION    AC144988.2  GI:32490606
KEYWORDS  HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE    Gorilla gorilla gorilla (lowland gorilla)
ORGANISM  Gorilla gorilla gorilla

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Gorilla.
1 (bases 1 to 222085)
Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Carriaga,K.,
Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X.,
Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P.,
Hurle,B., Idol,J.R., Karlins,E., Kwong,P., Laric,P., Lee-Lin,S.-Q.,
Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C.,
Maskeri,B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,M.E.,
Prasad,A., Reddix-Dugue,N., Schandler,K., Schueler,M.G., Shah,K.,
Sison,C., Stantripp,S., Thomas,J.W., Thomas,P.J., Tsipouri,V.,
Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 222085)
Green,E.D.
Direct Submission
Submitted (28-MAY-2003) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 222085)
Green,E.D.
Direct Submission
Submitted (10-JUL-2003) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
On Jul 10, 2003 this sequence version replaced gi:31088406.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: enc
Center Clone name: 209M14

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 220187 bases at least Q40
Consensus quality: 220809 bases at least Q30
Consensus quality: 221225 bases at least Q20
Insert size: 203000; agarose-fp
Insert size: 221385; sum-of-contigs
Quality coverage: 10.43x in Q20 bases; agarose-fp
Quality coverage: 9.56x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 2032: contig of 2032 bp in length

* 2033 2132: gap of unknown length

* 2133 40439: contig of 38307 bp in length

* 40440 40539: gap of unknown length

```
* 40540 59564: contig of 19025 bp in length
* 59565 59664: gap of unknown length
* 59665 93277: contig of 33613 bp in length
* 93278 93377: gap of unknown length
* 93378 154967: contig of 61589 bp in length
* 154967 155067: gap of unknown length
* 155067 168650: contig of 13584 bp in length
* 168651 168750: gap of unknown length
* 168751 185207: contig of 16457 bp in length
* 185208 185307: gap of unknown length
* 185308 222085: contig of 36778 bp in length.
FEATURES
    source
        1..222085
            /organism="Gorilla gorilla gorilla"
            /mol_type="genomic DNA"
            /sub_species="gorilla"
            /db_xref="taxon:9595"
            /clone="CH255-209M14"
            /clone_lib="CH255"
        1..2032
            /note="assembly_fragment
            missing T7 clone end on 5' end of insert"
            2033..2132
            /estimated_length=unknown
            /note="assembly_fragment"
            2133..40439
            /note="assembly_fragment"
            40440..40539
            /estimated_length=unknown
            /note="assembly_fragment"
            40540..59564
            /note="assembly_fragment"
            59565..59664
            /estimated_length=unknown
            59665..93277
            /note="assembly_fragment"
            93278..93377
            /estimated_length=unknown
            /note="assembly_fragment"
            93378..154966
            /note="assembly_fragment"
            154967..155066
            /estimated_length=unknown
            155067..168650
            /note="assembly_fragment"
            168651..168750
            /estimated_length=unknown
            168751..185207
            /note="assembly_fragment"
            185208..185307
            /estimated_length=unknown
            185308..222085
            /note="assembly_fragment
            clone_end:SP6
            vector_side:right"
    ORIGIN
        Query Match          90.9%; Score 80; DB 14; Length 222085;
        Best Local Similarity 94.3%; Pred No. 1.5e-16;
        Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 60
Db 116410 CGGAGCCGAGTCCTGATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 116351

Qy 61 AAATGACGAGTGTCTTCTGGTACAAAAA 89
Db 116350 AAATGACGAGTGTCTTCTGGTACAAAAA 116323

RESULT 12
BC006794
LOCUS
DEFINITION Homo sapiens interferon induced transmembrane protein 3 (1-8U),
mRNA (cdna clone MGC:5225 IMAGE:2986145), complete cds.
ACCESSION BC006794
```

```
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
CONSRMT
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
FEATURES
    source
        1..637
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="MGC:5225 IMAGE:2986145"
            /tissue_type="Cervix, carcinoma"
            /clone_lib="NIH_MGC_12"
            /lab_host="DH10B"
            /note="Vector: pCMV-SPORT6"
        1..637
            /gene="IFITM3"
            /note="synonym: 1-8U"
            /db_xref="GeneID:10410"
            /db_xref="MIM:605579"
            60..461
            /gene="IFITM3"
            /codon_start=1
    gene
    CDS
```

BC006794.1 GI:13905021
MGC.
Homo sapiens (human)
Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 637)
Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD,
Collins FS, Wagner L, Shenmen CW, Schuler GD, Altschul SF, Zeeberg
B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T,
Max SI, Wang J, Haieh F, Diatchenko L, Marusina K, Farmer AA, Rubin
GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL,
Scheetz TE, Brownstein MJ, Uedin TB, Toshiyuki S, Carninci P,
Prange C, Rana SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy
SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH,
Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW,
Villalon DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J,
Helton E, Kettaman M, Madan A, Rodrigues S, Sanchez A, Whiting M,
Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakealey RW,
Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, U,
Smailus DE, Schnerch A, Schein JE, Jones SJ and Marra MA.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 637)
Director MGC Project.
Direct Submission
Submitted (27-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAX Plate: 3 Row: j Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 11995467.
Location/Qualifiers
1..637
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:5225 IMAGE:2986145"
/tissue_type="Cervix, carcinoma"
/clone_lib="NIH_MGC_12"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..637
/gene="IFITM3"
/note="synonym: 1-8U"
/db_xref="GeneID:10410"
/db_xref="MIM:605579"
60..461
/gene="IFITM3"
/codon_start=1

```
/product="interferon-induced transmembrane protein 3
(1-8U)"
/protein_id="AAH06794.1"
/db_xref="GI:13905022"
/db_xref="GeneID:10410"
/db_xref="MIM:605579"
/translation="MNHTVQTFSPVNSGQPPNYEMLKEHEVAVLGAPHNPAPPTST
VHIRSETSPDHVWVLSLFTLFWNPCCLGFIAPFAYSVKSRDRKMGVDVTGAQAYAST
AKCLNIWALILGILMTILLIIVPLIFQYG"

ORIGIN
Query Match      89.1%; Score 78.4; DB 8; Length 637;
Best Local Similarity 93.2%; Pred. No. 8.1e-16;
Matches 82; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGCGATTCAAT 60
|||||
Db 544 CGGAGCCGAGTCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGCGATTCAAT 603
|||||

QY 61 AAGTGCACGTGTTCTTGGTACAAAA 88
|||||
Db 604 AAGTGCACGTGTTCTGGAATAAAAA 631
|||||

RESULT 13
AX281855 LOCUS AX281855 695 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 264 from Patent WO0177389.
ACCESSION AX281855
VERSION AX281855.1 GI:16609106
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Shiffman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
Mikita,T. and Tai,J
TITLE Genes expressed in foam cell differentiation
JOURNAL Patent: WO 0177389-A 264 18-OCT-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1..695
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 444902.6"

ORIGIN
Query Match      89.1%; Score 78.4; DB 6; Length 695;
Best Local Similarity 93.2%; Pred. No. 8.1e-16;
Matches 82; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGCGATTCAAT 60
|||||
Db 585 CGGAGCCGAGTCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGCGATTCAAT 644
|||||

QY 61 AAGTGCACGTGTTCTTGGTACAAAA 88
|||||
Db 645 AAGTGCACGTGTTCTGGAATAAAAA 672
|||||

RESULT 14
BC008417 LOCUS BC008417 749 bp mRNA linear PRI 29-JUN-2004
DEFINITION Homo sapiens interferon induced transmembrane protein 3
mRNA (cDNA clone MGC:14565 IMAGE:4075453), complete cds.
ACCESSION BC008417
VERSION BC008417.1 GI:14250037
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

REFERENCE
AUTHORS

1 (bases 1 to 749)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Udín,T.B., Toshiyuki,S.,
Carlinici,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Sahej,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Green,E.D.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Schmutz,J.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalek,U., Smalil,D.E.,
Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 749)

Strausberg,R.

Direct Submission

Submitted (25-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: CLONTECH Laboratories, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IPAL Plate: 21 Row: 0 Column: 15

This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.

FEATURES

source

1..749

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:14565 IMAGE:4075453"

/tissue_type="Brain, glioblastoma"

/clone_lib="NIH MGC_57"

/lab_host="DH10B"

/note="Vector: pDNR-LIB"

1..749

/gene="IFITM3"

/note="synonym: 1-8U"

/db_xref="GeneID:10410"

/db_xref="MIM:605579"

45..446

/gene="IFITM3"

/codon_start=1

/product="interferon-induced transmembrane protein 3
(1-8U)"

/protein_id="AAH08417.1"

gene

CDS

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 138.343 Seconds
(without alignments)
4239.411 Million cell updates/sec

Title: US-09-980-046B-5
Perfect score: 88
Sequence: 1 cggagtcgagtcctgtatca.....cgtgttcttctgtacaaaaa 88

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_21.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2003ds.*
12: geneseqn2004as.*
13: geneseqn2004bs.*
14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	88	4 AAC89736	Aac89736 Human gas
2	80	90.9	467	4 Aal07889	Aal07889 Human bre
3	80	90.9	483	9 ACH33912	Ach33912 Human end
4	80	90.9	536	4 Aal25635	Aal25635 Human bre
5	80	90.9	541	4 Aal18030	Aal18030 Human bre
6	80	90.9	561	10 ADK61256	Adk61256 Ovarian c
7	80	90.9	623	5 AAs75777	Aas75777 DNA encod
8	80	90.9	642	2 AAZ07228	Aaz07228 Human lun
9	80	90.9	642	3 AAC79118	Aac79118 Human lun
10	80	90.9	642	4 AAD23193	Aad23193 Human lun
11	80	90.9	642	10 ADD66427	Add66427 Human lun
12	80	90.9	642	10 ADE87681	Ade87681 Human lun
13	80	90.9	993	11 ACN88677	Acn88677 Breast ca
14	78.4	89.1	682	10 ADJ56210	Adj56210 Human cdn
15	78.4	89.1	695	6 AAs95009	Aas95009 Human DNA
16	76.8	87.3	451	2 AAZ41347	Aaz41347 Human nor
17	76	86.4	1408	12 ADP83378	Adp83378 Breast sp
18	76	86.4	1889	12 ADP83377	Adp83377 Breast sp
19	75.8	86.1	118	12 ADL83354	Adl83354 Human ost

20	75.8	86.1	167	2	AAT21450	Aat21450 Human gen
21	75.8	86.1	594	13	ADU14728	Adu14728 Solid tum
c 22	75.8	86.1	668	4	AAL26695	Aal26695 Human bre
23	75.8	86.1	808	3	AAA46668	Aaa46668 cDNA of I
24	75.8	86.1	808	8	ACC42318	Acc42318 Human MAP
25	75.8	86.1	808	10	ADK61255	Adk61255 Ovarian c
26	75.8	86.1	808	11	ADI31953	Adi31953 Human CDN
27	75.8	86.1	808	13	ADS84020	Ads84020 Human lym
28	75.8	86.1	808	14	ADX08047	Adx08047 Cyclin-de
29	75.8	86.1	808	14	ADY20076	Ady20076 DNA encod
30	75.8	86.1	808	14	ADY86641	Ady86641 Human int
31	75.8	86.1	808	14	AEA23687	Aea23687 Human PRO
32	75.4	85.7	606	8	ACC42317	Acc42317 Human MAP
33	75.4	85.7	621	8	ACC42319	Acc42319 Human MAP
34	74.8	85.0	545	10	ADK61257	Adk61257 Ovarian c
c 35	74.4	84.5	1933	1	AAN90120	Aan90120 Human lip
36	71	80.7	600	13	ADT50791	Adt50791 Cancer re
37	70.6	80.2	627	5	AAS75776	Aas75776 DNA encod
38	70.6	80.2	678	12	ADL07643	Adl07643 Human 1-8
39	70.6	80.2	1016	5	AAS83009	Aas83009 DNA encod
c 40	70.4	80.0	283	4	AAL20125	Aal20125 Human bre
41	70.2	79.8	761	6	ABQ56182	Abq56182 Human ova
42	69.6	79.1	407	3	AAF18371	Aaf18371 Lung canc
c 43	69.4	78.9	245	11	ACN90202	Acn90202 Breast ca
44	69	78.4	762	5	AAS87704	Aas87704 DNA encod
45	68.8	78.2	148834	6	ABK83570	Abk83570 Human CDN

ALIGNMENTS

RESULT 1
AAC89736
ID AAC89736 standard; cDNA; 88 BP.
XX
AC AAC89736;
XX
12-MAR-2001 (first entry)
XX
Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 5.
XX
Human; cytostatic; immunomodulator; immunostimulant; vulnery;
XX
anti-inflammatory; neuroprotective; antibacterial; gene therapy;
XX
gastrointestinal inflammation; immune system disorder; genetic disorder;
XX
cancer; autoimmune disorder; infection; wound healing; ss.
XX
Homo sapiens.
OS
WO200073324-A2.
FN
XX
07-DEC-2000.
PD
XX
01-JUN-2000; 2000WO-US015191.
PF
XX
01-JUN-1999; 99US-0137058P.
PR
XX
(DIGI-) DIGITAL GENE TECHNOLOGIES INC.
PA
XX
Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Hasel KW;
PI
XX
WPI; 2001-061508/07.
DR
XX
New polynucleotides and polypeptides, useful in gene therapy and in
XX
diagnosing a pathological condition, e.g. for modulating gene expression
PT
in gastrointestinal inflammation, or for treating cancers or genetic
PT
disorders.
PS
XX
Claim 1; Page 85; 108pp; English.
CC
The present sequence is one of a number of isolated human polynucleotides
CC
which are useful in gene therapy, and for diagnosing a pathological
CC
condition or a susceptibility to it. In particular, the polynucleotides
CC
are useful for modulating gene expression in gastrointestinal

CC inflammation. The polynucleotides are useful for chromosome
 CC identification, controlling gene expression through triple helix
 CC formation or antisense DNA or RNA, or identifying individuals from minute
 CC biological samples using DNA-based identification techniques. The
 CC polynucleotides can also be used as an alternative to restriction
 CC fragment length polymorphism (RFLP), by determining the actual base-by-
 CC base DNA sequences of selected portions of an individual's genome. The
 CC polynucleotides may also be used as molecular weight markers on Southern
 CC gels, as diagnostic probes for the presence of a specific mRNA, as a
 CC probe to subtract-out known sequences in the process of discovering novel
 CC polynucleotides, or as an antigen to elicit an immune response. The
 CC polynucleotides are useful in diagnostic procedures to detect a disorder.
 CC The polynucleotides and polypeptides are useful for preventing, treating
 CC or ameliorating immune system disorders, genetic disorders, cancers, some
 CC autoimmune disorders, or infections. The polynucleotides and polypeptides
 CC are also useful for differentiating, proliferating or attracting cells,
 CC leading to the regeneration of tissues, especially in wounds or burns.
 CC The polypeptides and polynucleotides may also be used as a food additive
 CC or preservative

XX
 SQ Sequence 88 BP; 23 A; 22 C; 16 G; 27 T; 0 U; 0 Other;

Query Match 100.0%; Score 88; DB 4; Length 88;
 Best Local Similarity 100.0%; Pred. No. 5.6e-22;
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 60
 Db |||||
 Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 60
 Db |||||
 Qy 61 AAAGTGCACGTGTTCTTCTGGTACAAAAA 88
 Db |||||
 Qy 61 AAAGTGCACGTGTTCTTCTGGTACAAAAA 88

RESULT 2

AAAL07889
 ID AAL07889 standard; cDNA; 467 BP.

AC AAL07889;

XX 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 346.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US000798.

XX 14-JAN-2000; 2000US-0176077P.

XX 14-MAR-2000; 2000US-0189167P.

XX 24-MAR-2000; 2000US-0192099P.

XX 29-MAR-2000; 2000US-0193480P.

XX 15-MAY-2000; 2000US-0205230P.

XX 09-JUN-2000; 2000US-0211315P.

XX 25-JUL-2000; 2000US-0220534P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer.

XX Claim 1; Page 153; 3695pp; English.

CC The invention relates to human breast cancer expressed polynucleotides
 CC (AAL07544-AAL26789) and methods of assessing whether a patient is
 CC afflicted with breast cancer by examining the correlation between the
 CC expression of certain markers and the cancerous state of breast cells.
 CC The polynucleotides and encoded polypeptides are potential markers for
 CC detecting, diagnosing, monitoring, characterizing treating and
 CC potentially preventing breast cancer. The polynucleotides and encoded
 CC polypeptides are also useful for isolating compounds with cytostatic
 CC activity

XX
 SQ Sequence 467 BP; 153 A; 112 C; 71 G; 131 T; 0 U; 0 Other;

Query Match 90.9%; Score 80; DB 4; Length 467;
 Best Local Similarity 94.3%; Pred. No. 7.2e-19;
 Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 60
 Db |||||
 Qy 67 CGGAGCGGAGTCCTGTATCAGCCCTTTATCCTCACAGCTTTTCTACAATGGCATTCAAT 126
 Db |||||
 Qy 61 AAAGTGCACGTGTTCTTCTGGTACAAAAA 88
 Db |||||
 Qy 127 AAAGTGCACGTGTTCTTCTGGGAGAAAAA 154

RESULT 3

ACH33912
 ID ACH33912 standard; cDNA; 483 BP.

XX ACH33912;

XX 13-OCT-2003 (first entry)

XX Human endothelial cell cDNA #2045.

XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
 XX genome mapping; biodiversity; genetic disorder.

XX Homo sapiens.

XX US2003073623-A1.

XX 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

XX (DRMA/) DRMANAC R T.

XX (LABA/) LABAT I.

XX (STAC/) STACHE-CRAIN B.

XX (DICK/) DICKSON M C.

XX (JONE/) JONES L W.

XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/SB.

XX New polynucleotide sequences obtained from various cDNA libraries, useful
 XX as hybridization probes, as oligomers for PCR, for chromosome and gene
 XX mapping, in the recombinant production of protein, or in generating
 XX antisense DNA or RNA.

XX Claim 1; SEQ ID NO 21124; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of
 XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
 XX determined by the technique of SBH (sequencing by hybridisation). Also
 XX included is a purified polypeptide comprising a sequence corresponding to
 XX a reading frame of the novel polynucleotide. The nucleic acid sequences
 XX are useful in diagnostics as expressed sequence tags (EST) for
 XX identifying expressed genes or for physical mapping of the human genome,
 XX in forensics, in assessing biodiversities, or in identifying mutations

CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030073623

XX SQ Sequence 483 BP; 121 A; 138 C; 112 G; 110 T; 0 U; 2 Other;
Query Match 90.9%; Score 80; DB 9; Length 483;
Best Local Similarity 94.3%; Pred. No. 7.3e-19;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCAGCGCTTTCTACAAATGGCATTCAAT 60
Db 217 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCAGCGCTTTCTACAAATGGCATTCAAT 276

Qy 61 AAGTGACAGTGTTCTTGGTACAAAA 88
Db 277 AAGTGACAGTGTTCTTGGTACAAAA 304

RESULT 4
AAL25635
ID AAL25635 standard; cDNA; 536 BP.

XX AC AAL25635;

XX DT 07-DEC-2001 (first entry)

XX DE Human breast cancer expressed polynucleotide 18092.

XX KW Human; breast cancer; cell marker; cytostatic; ss.

XX OS Homo sapiens.

XX FN WO200151628-A2.

XX PD 19-JUL-2001.

XX PF 10-JAN-2001; 2001WO-US000798.

XX PR 14-JAN-2000; 2000US-0176077P.

XX PR 14-MAR-2000; 2000US-0189167P.

XX PR 24-MAR-2000; 2000US-0192099P.

XX PR 29-MAR-2000; 2000US-0193480P.

XX PR 15-MAY-2000; 2000US-0205230P.

XX PR 09-JUN-2000; 2000US-0211315P.

XX PR 25-JUL-2000; 2000US-0220534P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Lillie J, Xu Y, Wang Y, Steinmann K;

XX XX WPI; 2001-451856/48.

XX DR New peptide useful as a marker for the diagnosis of breast cancer.

XX PS Claim 1; Page 3344; 3695pp; English.

XX CC The invention relates to human breast cancer expressed polynucleotides
(AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
XX activity

SQ Sequence 536 BP; 193 A; 110 C; 65 G; 168 T; 0 U; 0 Other;
Query Match 90.9%; Score 80; DB 4; Length 536;
Best Local Similarity 94.3%; Pred. No. 7.5e-19;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCAGCGCTTTCTACAAATGGCATTCAAT 60
Db 30 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCAGCGCTTTCTACAAATGGCATTCAAT 89

Qy 61 AAGTGACAGTGTTCTTGGTACAAAA 88
Db 90 AAGTGACAGTGTTCTTGGTACAAAA 117

RESULT 5

AAL18030
ID AAL18030 standard; cDNA; 541 BP.

XX AC AAL18030;

XX DT 07-DEC-2001 (first entry)

XX DE Human breast cancer expressed polynucleotide 10487.

XX KW Human; breast cancer; cell marker; cytostatic; ss.

XX OS Homo sapiens.

XX FN WO200151628-A2.

XX PD 19-JUL-2001.

XX PF 10-JAN-2001; 2001WO-US000798.

XX PR 14-JAN-2000; 2000US-0176077P.

XX PR 14-MAR-2000; 2000US-0189167P.

XX PR 24-MAR-2000; 2000US-0192099P.

XX PR 29-MAR-2000; 2000US-0193480P.

XX PR 15-MAY-2000; 2000US-0205230P.

XX PR 09-JUN-2000; 2000US-0211315P.

XX PR 25-JUL-2000; 2000US-0220534P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Lillie J, Xu Y, Wang Y, Steinmann K;

XX XX WPI; 2001-451856/48.

XX DR New peptide useful as a marker for the diagnosis of breast cancer.

XX PS Claim 1; Page 1874; 3695pp; English.

XX CC The invention relates to human breast cancer expressed polynucleotides
(AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
XX activity

SQ Sequence 541 BP; 194 A; 110 C; 68 G; 169 T; 0 U; 0 Other;

Query Match 90.9%; Score 80; DB 4; Length 541;

Best Local Similarity 94.3%; Pred. No. 7.5e-19;

Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCAGCGCTTTCTACAAATGGCATTCAAT 60
Db 34 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCAGCGCTTTCTACAAATGGCATTCAAT 93

AAZ07228
ID AAZ07228 standard; cDNA; 642 BP.
AC AAZ07228;
XX
DT 13-OCT-1999 (first entry)
XX
DE Human lung tumour protein SALT-T8 5' cDNA sequence.
XX
KW Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine;
KW immunotherapy; detection; inhibition; ss.
XX
OS Homo sapiens.
XX
XX WO9938973-A2.
XX
XX 05-AUG-1999.
XX
XX 26-JAN-1999; 99WO-US001642.
XX
XX 28-JAN-1998; 98US-00015022.
XX
XX 28-JAN-1998; 98US-00015029.
XX
XX 18-MAR-1998; 98US-00040828.
XX
XX 18-MAR-1998; 98US-00040831.
XX
XX 23-JUL-1998; 98US-00122191.
XX
XX 23-JUL-1998; 98US-00122192.
XX
XX 22-DEC-1998; 98US-00219245.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Lodes MJ, Frudakis TN, Mohamath R;
XX
XX WPI; 1999-479187/40.
XX
XX P-PSDB; AAY29544.
XX
XX Lung tumor specific polynucleotides for inhibiting the development of
XX lung cancer.
XX
XX Claim 13; Page 122; 171pp; English.
XX
XX The present invention describes lung tumour specific polynucleotides and
XX tumour antigens. AAZ07144 to AAZ07246 and AAZ08301 to AAZ08325 represent
XX specifically claimed polynucleotides, and AAY29486 to AAY29571 represent
XX amino acid sequences from the present invention. The lung tumour specific
XX polynucleotides and polypeptides can be used in pharmaceutical
XX compositions and vaccines to inhibit the development of lung cancer. They
XX can also be used to detect lung cancer in a patient. Probes and
XX antibodies derived from the lung tumour sequences are useful in detection
XX of lung cancer
XX
XX Sequence 642 BP; 142 A; 215 C; 147 G; 138 T; 0 U; 0 Other;
SQ
Query Match 90.9%; Score 80; DB 2; Length 642;
Best Local Similarity 94.3%; Pred. NO. 7.9e-19;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAACGCTTTCTACAATGGCATTCAAT 60
Db 535 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCACAACGCTTTCTACAATGGCATTCAAT 594
Qy 61 AAAGTGCACGTGTTTCTTGGTACAAAAA 88
Db 595 AAAGTGCACGTGTTTCTTGGTACAAAAA 622
Query Match 90.9%; Score 80; DB 2; Length 642;
Best Local Similarity 94.3%; Pred. NO. 7.9e-19;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAACGCTTTCTACAATGGCATTCAAT 60
Db 535 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCACAACGCTTTCTACAATGGCATTCAAT 594
Qy 61 AAAGTGCACGTGTTTCTTGGTACAAAAA 88
Db 595 AAAGTGCACGTGTTTCTTGGTACAAAAA 622
RESULT 9
AAC79118
ID AAC79118 standard; cDNA; 642 BP.
XX
XX AAC79118;
AC
XX
XX 05-FEB-2001 (first entry)
DT
XX

DE Human lung tumour-specific cDNA #71.
XX
KW Lung tumour protein; lung cancer; cytostatic; vaccine; ss.
XX
OS Homo sapiens.
XX
XX WO200060077-A2.
XX
XX 12-OCT-2000.
XX
XX 30-MAR-2000; 2000WO-US008560.
XX
XX 02-APR-1999; 99US-00285323.
XX
XX 09-AUG-1999; 99US-00370838.
XX
XX 30-DEC-1999; 99US-00476235.
XX
XX 03-MAR-2000; 2000US-00518809.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Lodes MJ, Mohamath R, Secrist H;
XX
XX WPI; 2000-638466/61.
XX
XX P-PSDB; AAB44456.
XX
XX Novel lung tumor polypeptides and polynucleotides, useful for detecting,
XX monitoring or treating cancer, especially lung cancer.
XX
XX Claim 1; Page 139; 243pp; English.
XX
XX The present sequence is given in a specification relating to compounds
XX for therapy and diagnosis of lung cancer. Polypeptides comprising at
XX least an immunogenic part of a lung tumour protein are disclosed. The
XX polypeptides are useful for inhibiting the development of cancer,
XX especially lung cancer. Samples of T cells expressing the polypeptides
XX may be used to inhibit the development of cancer. The polypeptides are
XX also useful for detecting and monitoring the progression of cancer,
XX especially lung cancer
XX
XX Sequence 642 BP; 142 A; 215 C; 147 G; 138 T; 0 U; 0 Other;
SQ
Query Match 90.9%; Score 80; DB 3; Length 642;
Best Local Similarity 94.3%; Pred. NO. 7.9e-19;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAACGCTTTCTACAATGGCATTCAAT 60
Db 535 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCACAACGCTTTCTACAATGGCATTCAAT 594
Qy 61 AAAGTGCACGTGTTTCTTGGTACAAAAA 88
Db 595 AAAGTGCACGTGTTTCTTGGTACAAAAA 622
RESULT 10
AAD23193
ID AAD23193 standard; cDNA; 642 BP.
XX
XX AAD23193;
AC
XX
XX 26-FEB-2002 (first entry)
DT
XX
XX Human lung tumour-specific protein SALT-T8 cDNA.
XX
XX Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
KW anticense-therapy; vaccine; immune response; lung cancer; SALT-T8; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 51..452
XX /*tag= a
XX /product= "Human SALT-T8 protein"
XX

```
PN WO200172295-A2.
XX
PD 04-OCT-2001.
XX
XX 28-MAR-2001; 2001WO-US009991.
XX
XX 29-MAR-2000; 2000US-00538037.
XX
XX 05-JUN-2000; 2000US-00588937.
XX
XX 18-AUG-2000; 2000US-00640878.
XX
XX 22-SEP-2000; 2000US-0234517P.
XX
XX 01-NOV-2000; 2000US-00704512.
XX
XX 14-DEC-2000; 2000US-00738973.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY;
PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX
XX WPI; 2001-639201/73.
XX
XX P-PSDB; AAE13797.
XX
XX New human lung-specific polynucleotides and polypeptides for the
PT diagnosis and treatment of disease e.g. lung cancer.
XX
XX Example 6; Page 193; 378pp; English.
XX
XX The invention relates to isolated lung tumour-specific proteins and their
CC corresponding cDNA molecules. Lung tumour-specific proteins and their
CC antigen-presenting cells are useful for stimulating and/or expanding T
CC cells specific for a tumour protein, and for inhibiting the development
CC of cancer. The invention also relates to a composition useful for
CC stimulating an immune response, and for treating cancer. The lung tumour
CC specific oligonucleotide is useful in gene therapy and for diagnosis,
CC detection and treatment of lung cancer. The present sequence is a cDNA
CC encoding human lung tumour-specific protein
XX
XX Sequence 642 BP; 142 A; 215 C; 147 G; 138 T; 0 U; 0 Other;
SQ
Query Match 90.9%; Score 80; DB 4; Length 642;
Best Local Similarity 94.3%; Pred. No. 7.9e-19;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 CGAGTCGAGTCCTGTATCAGCCCTTTATCTCCTCACAGCTTTTCTACAATGGCATTCAAT 60
DB 535 CGAGCGCGAGTCCTGTATCAGCCCTTTATCTCCTCACAGCTTTTCTACAATGGCATTCAAT 594
QY 61 AAAGTGCACGTGTTTCTTGGTGACAAAAA 88
DB 595 AAAGTGCACGTGTTTCTTGGTGAAAAAAA 622
RESULT 11
ADD66427
ID ADD66427 standard; cDNA; 642 BP.
XX
XX ADD66427;
XX
XX 15-JAN-2004 (first entry)
XX
XX Human lung tumour-specific related cDNA, SEQ ID No 119.
XX
XX expression control; cancer; T cell; tumour; immune; cytostatic; vaccine;
XX human; lung tumour-specific; ss.
XX
XX Homo sapiens.
XX
XX WO200292001-A2.
XX
XX 21-NOV-2002.
XX
XX 10-MAY-2002; 2002WO-US014975.
XX
XX 11-MAY-2001; 2001US-00854133.
XX
XX (CORI-) CORIXA CORP.
XX
XX Lodes MJ, Wang T, Fan L, Algate PA, Mcneill PD;
PI
XX WPI; 2003-120592/11.
XX
XX New polynucleotide and polypeptide, useful for preparing a composition
PT for diagnosing, treating or preventing cancer.
XX
XX Disclosure; SEQ ID NO 119; 494pp; English.
XX
XX The invention relates to a novel isolated polynucleotide comprising one
CC of 32 47-6080 base pair sequences, given in the specification, or their
CC complements or degenerate variants, at least 20 contiguous residues of a
CC sequence in, or having at least 75 or 90 % identity with the isolated
CC polynucleotide, or that hybridise with the polynucleotide. The invention
CC further comprises: an isolated polypeptide; an expression vector
CC comprising the polynucleotide operably linked to an expression control
CC sequence; a host cell transformed or transfected with the expression
CC vector; an isolated antibody or its antigen-binding fragment that
CC specifically binds to the polypeptide; a method for detecting the
CC presence of a cancer in a patient; a fusion protein comprising the
CC polypeptide; an oligonucleotide that hybridises to the isolated
CC polynucleotide under moderately stringent conditions; a method for
CC stimulating and/or expanding T cells specific for a tumour protein; an
CC isolated T cell population; a composition comprising a first component
CC consisting of carriers and immunostimulants and a second component; a
CC method for stimulating an immune response in a patient; a method for
CC treating cancer in a patient; a method for determining cancer in a
CC patient; a diagnostic kit comprising at least one oligonucleotide or
CC antibody and a detection reagent comprising a reporter group; and a
CC method for inhibiting the development of cancer in a patient. The
CC compositions of the invention have cytostatic activity and can be used to
CC create a vaccine. The isolated polynucleotide is useful for preparing a
CC composition for diagnosing, treating or preventing cancer. This
CC polynucleotide sequence represents a human lung tumour-specific cDNA
CC sequence relating to the invention.
XX
XX Sequence 642 BP; 142 A; 215 C; 147 G; 138 T; 0 U; 0 Other;
SQ
Query Match 90.9%; Score 80; DB 10; Length 642;
Best Local Similarity 94.3%; Pred. No. 7.9e-19;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 CGAGTCGAGTCCTGTATCAGCCCTTTATCTCCTCACAGCTTTTCTACAATGGCATTCAAT 60
DB 535 CGAGCGCGAGTCCTGTATCAGCCCTTTATCTCCTCACAGCTTTTCTACAATGGCATTCAAT 594
QY 61 AAAGTGCACGTGTTTCTTGGTGACAAAAA 88
DB 595 AAAGTGCACGTGTTTCTTGGTGAAAAAAA 622
RESULT 12
ADE87681
ID ADE87681 standard; cDNA; 642 BP.
XX
XX ADE87681;
XX
XX 29-JAN-2004 (first entry)
XX
XX Human lung tumour antigen cDNA #71.
XX
XX Human; lung tumour antigen; gene; ss; cancer; lung cancer; CD4+; CD8+;
XX T cell; immune response; immunostimulant; cytostatic.
XX
XX Homo sapiens.
XX
XX US2003118599-A1.
XX
XX 26-JUN-2003.
XX
```



```
XX 23-FEB-2001; 2001US-0270784P.
PR (STUA/) STUART S G.
XX (NUCH/) NUCHTERN J G.
PA (PLON/) PLON S E.
PA (SHOH/) SHOHET J M.
XX
PI Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
XX
DR WPI; 2003-635698/60.
XX
PT New genes regulated by MYCN activation, useful in gene therapy, other
PT particularly for treating a subject with e.g. neuroblastoma or other
PT cancers, or for diagnosing, staging or monitoring the treatment of the
PT cancer.
XX
PS Claim 1; SEQ ID NO 16; 27pp; English.
XX
CC This invention relates to novel isolated cDNAs that are differentially
CC expressed in MYCN activated cells. Specifically, it refers to
CC polynucleotide sequences that exhibit differential expression patterns in
CC cells activated by the transactivator MYCN, where MYCN is a proto-
CC oncogene that is amplified in neuroblastoma cells and is common in small
CC cell lung cancers. The present invention describes these cDNA molecules
CC as useful for in hybridisation assays to detect expression of nucleic
CC acids (or complementary nucleic acids) in a present in a given sample, as
CC well as for screening assays by identifying molecules or compounds that
CC specifically bind the cDNA as a ligand and modulate function or activity.
CC Accordingly, these compositions exhibit cytostatic activity and can also
CC be used for gene therapy purposes. This polynucleotide sequence is a cDNA
CC that is differentially expressed in MYCN activated cells, given in an
CC exemplification of the invention. NOTE: This sequence does not appear in
CC the printed specification but has been obtained in electronic format from
CC the US Patent Office at
CC ftp.seqdata.uspto.gov/sequence.html?DocID=20030119009.
XX
SQ Sequence 682 BP; 152 A; 225 C; 161 G; 144 T; 0 U; 0 Other;

Query Match      89.1%; Score 78.4; DB 10; Length 682;
Best Local Similarity 93.2%; Pred. No. 3.1e-18;
Matches 82; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCAGCGCTTTTCTACAATGGCATTCAAT 60
DB 572 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCAGCGCTTTTCTACAATGGCATTCAAT 631
QY 61 AAAGTCACGCTGTTTCTTGGTACAAAA 88
DB 632 AAAGTCACGCTGTTTCTTGGTACAAAA 659

RESULT 15
AAS95009
ID AAS95009 standard; DNA; 695 BP.
XX
AC AAS95009;
XX
XX 14-FEB-2002 (first entry)
XX
DE Human DNA sequence #264 expressed during foam cell differentiation.
XX
KW Human; foam cell differentiation; atherosclerosis; cerebral stroke;
KW cardiovascular disorder; coronary artery disease; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN W0200177389-A2.
XX
PD 18-OCT-2001.
XX
XX 04-APR-2001; 2001WO-US011128.
XX

05-APR-2000; 2000US-0195106P.
PR (INCY-) INCYTE GENOMICS INC.
XX
PI Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;
PI Tai J;
XX
DR WPI; 2002-010925/01.
XX
PT Composition useful for diagnosis of conditions, disorders or diseases
PT associated with atherosclerosis, comprises several polynucleotides that
PT are differentially expressed in foam cell development.
XX
PS Claim 1; Page 303-304; 315pp; English.
XX
CC The present invention relates to the isolation of human polynucleotide
CC sequences that are differentially expressed during foam cell
CC differentiation. The polynucleotide sequences of the invention or a
CC composition comprising these polynucleotides are useful as a high
CC throughput method for detecting altered expression of one or more
CC polynucleotides in a sample. The polynucleotides can be used in the
CC diagnosis of disorders associated with foam cell development such as
CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
CC coronary artery disease. The polynucleotide sequences can also be used as
CC PCR primers and probes. The polynucleotides of the invention are also
CC useful in gene therapy. AAS94746-AAS95021 represent the human
CC polynucleotide sequences of the invention which are differentially
CC expressed during foam cell differentiation
XX
SQ Sequence 695 BP; 158 A; 226 C; 166 G; 145 T; 0 U; 0 Other;

Query Match      89.1%; Score 78.4; DB 6; Length 695;
Best Local Similarity 93.2%; Pred. No. 3.1e-18;
Matches 82; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCAGCGCTTTTCTACAATGGCATTCAAT 60
DB 585 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCAGCGCTTTTCTACAATGGCATTCAAT 644
QY 61 AAAGTCACGCTGTTTCTTGGTACAAAA 88
DB 645 AAAGTCACGCTGTTTCTTGGTACAAAA 672

Search completed: December 7, 2005, 00:03:04
Job time : 141.343 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:31:40 ; Search time 219.122 Seconds
(without alignments)
3320.997 Million cell updates/sec

Title: US-09-980-046B-5
Perfect score: 88
Sequence: 1 cggagtcgagtcctgtatca.....cgtgttcttggtaaaaaa 88

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:

- 1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	80	90.9	483	Sequence 21124, A
2	80	90.9	561	Sequence 426, App
3	80	90.9	623	Sequence 11581, A
4	80	90.9	642	Sequence 119, App
5	80	90.9	642	Sequence 119, App
6	80	90.9	642	Sequence 119, App
7	80	90.9	993	Sequence 9827, App
8	78.4	89.1	682	Sequence 16, Appl
9	78.4	89.1	695	Sequence 264, App
10	75.8	86.1	118	Sequence 15601, A
11	75.8	86.1	118	Sequence 15601, A
12	75.8	86.1	118	Sequence 15601, A
13	75.8	86.1	596	Sequence 3, Appl
14	75.8	86.1	808	Sequence 73, Appl
15	75.8	86.1	808	Sequence 55, Appl
16	75.8	86.1	808	Sequence 55, Appl
17	75.8	86.1	808	Sequence 1279, App
18	75.8	86.1	808	Sequence 31, Appl
19	75.8	86.1	808	Sequence 425, App
20	75.8	86.1	808	Sequence 2340, App
21	75.8	86.1	808	Sequence 7575, App
22	75.4	85.7	401	Sequence 40680, A
23	75.4	85.7	401	Sequence 40680, A

Sequence 71, Appl
Sequence 75, Appl
Sequence 427, App
Sequence 6787, App
Sequence 6787, App
Sequence 279728, App
Sequence 279728, App
Sequence 279728, App
Sequence 105978, App
Sequence 105978, App
Sequence 18813, A
Sequence 18813, A
Sequence 18815, A
Sequence 18815, A
Sequence 18816, A
Sequence 18816, A
Sequence 18813, A
Sequence 18813, A
Sequence 18815, A
Sequence 18815, A
Sequence 41318, A
Sequence 41318, A
Sequence 41318, A
Sequence 11580, A

24 75.4 85.7 606 5 US-10-197-666A-71
25 75.4 85.7 621 5 US-10-197-666A-75
26 74.8 85.0 545 5 US-10-505-680-427
27 72.4 82.3 346 7 US-10-242-535A-6787
28 72.4 82.3 346 7 US-10-085-783A-6787
29 72 81.8 546 5 US-10-027-632-279728
30 72 81.8 546 5 US-10-027-632-279728
31 72 81.8 546 6 US-10-027-632-279728
32 72 81.8 546 6 US-10-027-632-279728
33 72 81.8 575 5 US-10-027-632-105978
34 72 81.8 575 6 US-10-027-632-105978
35 72 81.8 714 5 US-10-027-632-19813
36 72 81.8 714 5 US-10-027-632-19814
37 72 81.8 714 5 US-10-027-632-19815
38 72 81.8 714 5 US-10-027-632-19816
39 72 81.8 714 6 US-10-027-632-19813
40 72 81.8 714 6 US-10-027-632-19814
41 72 81.8 714 6 US-10-027-632-19815
42 72 81.8 714 6 US-10-027-632-19816
43 71.8 81.6 189 7 US-10-242-535A-41318
44 71.8 81.6 189 7 US-10-085-783A-41318
45 70.6 80.2 627 9 US-10-450-763-11580

ALIGNMENTS

RESULT 1
US-09-918-995-21124
; Sequence 21124, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq. Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21124
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)...(483)
; OTHER INFORMATION: n = A,T,C or G

Query Match 90.9%; Score 80; DB 3; Length 483;
Best Local Similarity 94.3%; Pred. No. 2.2e-18;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGATCAGCCCTTTCTCTACACGCTTTCTTACAAATGCGATTCAAT 60
Db 217 CGGAGTCGAGTCCTGATCAGCCCTTTCTCTACACGCTTTCTTACAAATGCGATTCAAT 276
Qy 61 AAAGTGCACGCTGTTCTTGGTACAAAAA 88
Db 277 AAAGTGCACGCTGTTCTTGGTACAAAAA 304

RESULT 2
US-10-505-680-426/c
; Sequence 426, Application US/10505680
; Publication No. US2005009592A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America as represented by the
; APPLICANT: Secretary of the Department of Health and Human Services
; APPLICANT: Jazaeri, Amir A.

```
; APPLICANT: Boyd, Jeff
; APPLICANT: Liu, Edison T.
; TITLE OF INVENTION: IDENTIFICATION OF OVARIAN CANCER TUMOR MARKERS AND THERAPEUTIC TA
; FILE REFERENCE: 4239-64452
; CURRENT APPLICATION NUMBER: US/10/505,680
; CURRENT FILING DATE: 2004-08-12
; PRIOR APPLICATION NUMBER: 60/357,031
; PRIOR FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 822
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 426
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-505-680-426

Query Match          90.9%; Score 80; DB 9; Length 561;
Best Local Similarity 94.3%; Pred. No. 2.3e-18;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 60
Db 88 CGGAGCGGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 29

Qy 61 AAAGTGCACGTGTTTCTTGGTACAAAAA 88
Db 28 AAAGTGCACGTGTTTCTGTTGTAATAAAAAA 1

RESULT 3
US-10-450-763-11581
; Sequence 11581, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 11581
; LENGTH: 623
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: SIMILAR
; LOCATION: (49)..(366)
; OTHER INFORMATION: 96% homologous to Homo sapiens 1-8U, accession number
; OTHER INFORMATION: X57352, Smith-Waterman Score=541.
US-10-450-763-11581

Query Match          90.9%; Score 80; DB 9; Length 623;
Best Local Similarity 94.3%; Pred. No. 2.4e-18;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 60
Db 535 CGGAGCGGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 594

Qy 61 AAAGTGCACGTGTTTCTTGGTACAAAAA 88
Db 595 AAAGTGCACGTGTTTCTGTTGTAATAAAAAA 622

RESULT 4
US-09-738-973-119
```

```
; Sequence 119, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 119
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-738-973-119

Query Match          90.9%; Score 80; DB 3; Length 642;
Best Local Similarity 94.3%; Pred. No. 2.4e-18;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 60
Db 535 CGGAGCGGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 594

Qy 61 AAAGTGCACGTGTTTCTTGGTACAAAAA 88
Db 595 AAAGTGCACGTGTTTCTGTTGTAATAAAAAA 622

RESULT 5
US-09-854-133-119
; Sequence 119, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 119
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-854-133-119

Query Match          90.9%; Score 80; DB 3; Length 642;
Best Local Similarity 94.3%; Pred. No. 2.4e-18;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 60
Db 535 CGGAGCGGAGTCCTGTATCAGCCCTTTATCCTCACAGCCTTTTCTACAATGGCATTCAAT 594

Qy 61 AAAGTGCACGTGTTTCTTGGTACAAAAA 88
```


Db 595 AAGTGCACGTGTTCTTGGTGAACAAA 622
|||||

RESULT 6

US-10-144-649A-119
; Sequence 119, Application US/10144649A
; Publication No. US20030118599A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Algate, Paul A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144,649A
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 119
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-144-649A-119

Query Match 90.9%; Score 80; DB 5; Length 642;
Best Local Similarity 94.3%; Pred. No. 2.4e-18;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGGCATTCAAT 60
Db 535 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGGCATTCAAT 594
|||||

Qy 61 AAGTGCACGTGTTCTTGGTACAAAAA 88
|||||

Db 595 AAGTGCACGTGTTCTTGGTGAACAAA 622
|||||

RESULT 7

US-10-198-846-9827
; Sequence 9827, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9827
; LENGTH: 993
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-9827

Query Match 90.9%; Score 80; DB 5; Length 993;
Best Local Similarity 94.3%; Pred. No. 2.8e-18;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGGCATTCAAT 60
Db 488 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGGCATTCAAT 547
|||||

Qy 61 AAGTGCACGTGTTCTTGGTACAAAAA 88
|||||
Db 548 AAGTGCACGTGTTCTTGGTGAACAAA 575
|||||

RESULT 8

US-10-084-817-16
; Sequence 16, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 1101453.2
US-10-084-817-16

Query Match 89.1%; Score 78.4; DB 5; Length 682;
Best Local Similarity 93.2%; Pred. No. 9.4e-18;
Matches 82; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGGCATTCAAT 60
Db 572 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGGCATTCAAT 631
|||||

Qy 61 AAGTGCACGTGTTCTTGGTACAAAAA 88
|||||

Db 632 AAGTGCACGTGTTCTTGGAAATAAAA 659
|||||

RESULT 9

US-10-240-965-264
; Sequence 264, Application US/10240965
; Publication No. US20030165924A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: SHIFFMAN, Dov
; APPLICANT: SOMOGYI, Roland
; APPLICANT: LAWN, Richard M.
; APPLICANT: SEILHAWER, Jeffrey J.
; APPLICANT: PORTER, Gordon J.
; APPLICANT: MIKITA, Thomas
; APPLICANT: TAI, Julie
; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
; FILE REFERENCE: PA-0025 PCT
; CURRENT APPLICATION NUMBER: US/10/240,965
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/195,106
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 276
; SOFTWARE: PERL Program
; SEQ ID NO 264
; LENGTH: 695
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030165924A1 444902.6
US-10-240-965-264

; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 9
; LENGTH: 596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (22)..(423)
US-10-115-831-9

Query Match 86.1%; Score 75.8; DB 6; Length 596;
Best Local Similarity 97.5%; Pred. No. 7.7e-17;
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCATTCAAT 60
|||
Db 506 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCATTCAAT 565
|||
Qy 61 AAAGTGCACGTGTTTCTTG 79
|||
Db 566 AAAGTGCACGTGTTTCTGG 584
|||

RESULT 14
US-10-197-666A-73
; Sequence 73, Application US/10197666A
; Publication No. US20030092037A1
; GENERAL INFORMATION:
; APPLICANT: ASAH KASEI KABUSIKI KAISYA
; TITLE OF INVENTION: E1k1 phosphorylation related gene
; FILE REFERENCE: PH-1548US
; CURRENT APPLICATION NUMBER: US/10/197,666A
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305,884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316,304
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350,027
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 73
; LENGTH: 808
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (238)..(636)
US-10-197-666A-73

Query Match 86.1%; Score 75.8; DB 5; Length 808;
Best Local Similarity 97.5%; Pred. No. 8.7e-17;
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCATTCAAT 60
|||
Db 722 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCATTCAAT 781
|||
Qy 61 AAAGTGCACGTGTTTCTTG 79
|||
Db 782 AAAGTGCACGTGTTTCTGG 800
|||

RESULT 15
US-10-133-937-55
; Sequence 55, Application US/10133937
; Publication No. US2003020278A1

; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING,
; TITLE OF INVENTION: DIAGNOSING, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND
; TITLE OF INVENTION: OTHER BIOLOGICAL STATES
; FILE REFERENCE: 11613.56US01
; CURRENT APPLICATION NUMBER: US/10/133,937
; CURRENT FILING DATE: 2002-11-04
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 55
; LENGTH: 808
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-133-937-55

Query Match 86.1%; Score 75.8; DB 6; Length 808;
Best Local Similarity 97.5%; Pred. No. 8.7e-17;
Matches 77; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCATTCAAT 60
|||
Db 722 CGGAGCCGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCATTCAAT 781
|||
Qy 61 AAAGTGCACGTGTTTCTTG 79
|||
Db 782 AAAGTGCACGTGTTTCTGG 800
|||

Search completed: December 6, 2005, 22:57:05
Job time : 220.122 secs

This Page Blank (uap16)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:34:11 ; Search time 60.0707 Seconds
(without alignments)
455.919 Million cell updates/sec

Title: US-09-980-046B-5
Perfect score: 88
Sequence: 1 cggagtcgagtcctgtatca.....cgtgtttcttggtacaaaaa 88

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New: *
1: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	41.2	46.8	182314	7	US-11-112-908-45
C 2	30	34.1	4116	6	US-10-750-185-38038
C 3	27.2	30.9	1033	6	US-10-750-185-57623
C 4	25.6	29.1	1668	6	US-10-750-185-50602
C 5	25.4	28.9	2224	6	US-10-750-185-53492
C 6	25.2	28.6	1642	6	US-10-750-185-41665
C 7	24.8	28.2	1998	6	US-10-750-185-45955
C 8	24.6	28.0	494	6	US-10-750-185-2846
C 9	24.2	27.5	2156	6	US-10-750-185-51326
C 10	24.2	27.5	2357	6	US-10-750-185-56680
C 11	24	27.3	3541	6	US-10-750-185-24928
C 12	23.8	27.0	1049	6	US-10-750-185-64427
C 13	23.8	27.0	2248	6	US-10-750-185-25851
C 14	23.8	27.0	2301	6	US-10-793-626-419
C 15	23.8	27.0	2744	6	US-10-793-626-4069
C 16	23.8	27.0	3127	6	US-10-793-626-3502
C 17	23.8	27.0	3170	6	US-10-793-626-3892
C 18	23.8	27.0	4094	6	US-10-793-626-3641
C 19	23.8	27.0	6968	6	US-10-793-626-759
C 20	23.6	26.8	1759	6	US-10-750-185-57531
C 21	23.4	26.6	1407	6	US-10-750-185-34679
C 22	23.4	26.6	1506	6	US-10-750-185-44563
C 23	23.4	26.6	1987	6	US-10-750-185-62265

ALIGNMENTS

RESULT 1

US-11-112-908-45/c
; Sequence 45, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 45
; LENGTH: 182314
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-45

Query Match 46.8%; Score 41.2; DB 7; Length 182314;
Best Local Similarity 74.3%; Pred. No. 3.2e-05;
Matches 52; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 19 CAGCCCTTATCTCCACACGCTTTTCTACATGCGATCAATGAAGTCGACGTGTTCTT 78
Db 27145 CAGCAGTTTATCTACATCACTTTTCTACATGTTTGTGATGTCAGGCGATGTTGTTG 27086
Qy 79 GGTACAAAAA 88
Db 27085 AAAAAAAA 27076

RESULT 2

US-10-750-185-38038
; Sequence 38038, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.

```
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 38038
; LENGTH: 4116
; TYPE: DNA
; ORGANISM: Bovine 19866881138467
US-10-750-185-38038

Query Match      34.1%; Score 30; DB 6; Length 4116;
Best Local Similarity 72.2%; Pred. No. 0.11;
Matches 39; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 22 CCCTTTATCTCCACACGCTTTTCTACAAATGGCATTCAATAAAGTGACGTTT 75
Db 1469 CCGTCTAATCTGATATGCTTTCTCCCAATGGCTTCTCTACAGTGAACAGGTTT 1522

RESULT 3
US-10-750-185-57623
; Sequence 57623, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 57623
; LENGTH: 1033
; TYPE: DNA
; ORGANISM: Bovine 19866881623743
US-10-750-185-57623

Query Match      30.9%; Score 27.2; DB 6; Length 1033;
Best Local Similarity 64.1%; Pred. No. 0.73;
Matches 41; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 22 CCCTTTATCTCCACACGCTTTTCTACAAATGGCATTCAATAAAGTGACGTTTCTTGGT 81
Db 177 CCCTTTAACCACACCTCTCTTTTAACTCTGCAATTCATTTCTGTCCATATTTTCAAGAT 236

Qy 82 ACAA 85
Db 237 CCTA 240

RESULT 4
US-10-750-185-50602
; Sequence 50602, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
```

```
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 50602
; LENGTH: 1668
; TYPE: DNA
; ORGANISM: Bovine 19866880988770
US-10-750-185-50602

Query Match      29.1%; Score 25.6; DB 6; Length 1668;
Best Local Similarity 70.8%; Pred. No. 3.3;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 41 TTTCTCAATGGCATTCAATAAAGTGACGTTTCTTGTGTAACAAAA 88
Db 222 TTTCTCAATGACATTCATTAAGTCCAGAGGCTGGAGGNAAAAAA 269

RESULT 5
US-10-750-185-53492/c
; Sequence 53492, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 53492
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Bovine 19866880905627
US-10-750-185-53492

Query Match      28.9%; Score 25.4; DB 6; Length 2224;
Best Local Similarity 64.4%; Pred. No. 4.3;
Matches 38; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 11 TCCTGTATCAGCCCTTTATCCTCACACGCTTTTCTCAATGGCATTCAATAAAGTGAC 69
Db 983 TCCTTTATCAGCTTTTATATATATATGATGATTTTAAATAAGCAATCCATAGAAGAAC 925

RESULT 6
US-10-750-185-41665/c
; Sequence 41665, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
```

; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41665
; LENGTH: 1642
; TYPE: DNA
; ORGANISM: Bovine 19866880829778
US-10-750-185-41665

Query Match 28.6%; Score 25.2; DB 6; Length 1642;
Best Local Similarity 62.9%; Pred. No. 4.6;
Matches 39; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 27 TATCCTCACGCTTTTCTACAAATGCGATTCAATAAGTGCACGTGTTCTTGTGTAACA 86
Db 137 TTTTCTATCCCTCTTTCTACTATAGCATTTCTCTCAAAAGCATTTCTTCTCTTCAATAA 78
Qy 87 AA 88
Db 77 AA 76

RESULT 7
US-10-750-185-45955/c
; Sequence 45955, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45955
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Bovine 19866880580954
US-10-750-185-45955

Query Match 28.2%; Score 24.8; DB 6; Length 1998;
Best Local Similarity 56.0%; Pred. No. 6.9;
Matches 47; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 5 GTGAGTCCTGTATCAGCCCTTTATCCTCACACGCTTTCTACATGCGATTGCAATTAAG 64
Db 1263 GTAGCTTCGTGAGCAGCAGGTTCATCTGTCTGCGCTTTAAAGCAATGCAAGAGC 1204
Qy 65 TGCAGCTGTTTCTTGGTACAAAA 88
Db 1203 TGTAGACCTCTTAGCTATATAA 1180

RESULT 8
US-10-750-185-2846
; Sequence 2846, Application US/10750185

; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2846
; LENGTH: 494
; TYPE: DNA
; ORGANISM: Bovine MMBT14110
US-10-750-185-2846

Query Match 28.0%; Score 24.6; DB 6; Length 494;
Best Local Similarity 57.0%; Pred. No. 5.1;
Matches 45; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 9 AGTCCTGTATCAGCCCTTTATCCTCACACGCTTTCTACAAATGCGATTGCAATTAAGTGA 68
Db 178 ATTGCTGTAGGACCTCTCACTCTCTCACGCTTAATTAGAAATGCAATAAACCAAAAGA 237
Qy 69 CGTGTCTTCTGGTACAAAA 87
Db 238 TTTTCTTTTCTCTGCTTAA 256

RESULT 9
US-10-750-185-51326/c
; Sequence 51326, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51326
; LENGTH: 2156
; TYPE: DNA
; ORGANISM: Bovine 19866880417584
US-10-750-185-51326

Query Match 27.5%; Score 24.2; DB 6; Length 2156;
Best Local Similarity 71.1%; Pred. No. 12;
Matches 32; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 33 CACAGCTTTTCTACAATGCGATTCAATAAGTGCAGTGTCTTCT 77
Db 1973 CACATTCTTTTCCATACAGGCTTTCTGTTAAGTCAGGTGAATCT 1929

RESULT 10
US-10-750-185-56680

; Sequence 56680, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 56680
; LENGTH: 2357
; TYPE: DNA
; ORGANISM: Bovine 19866880352383
US-10-750-185-56680

Query Match 27.5%; Score 24.2; DB 6; Length 2357;
Best Local Similarity 59.4%; Pred. No. 12; Mismatches 28; Indels 0; Gaps 0;
Matches 41; Conservative 0;

Qy 20 AGCCCTTTATCCTCACACGCTTTTCTACAATGGCAATCAATAAAGTGCACGTGTTCTTG 79
Db 1326 AGGCCTGGAATCAGGAGCGCTGGCACACGCGCACTCAAGNAAGCTCTTGTGTTGAAG 1385

Qy 80 GTACAAAAA 88
Db 1386 GAACAACA 1394

RESULT 11
US-10-750-185-24928
; Sequence 24928, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 24928
; LENGTH: 3541
; TYPE: DNA
; ORGANISM: Bovine 19866880691281
US-10-750-185-24928

Query Match 27.3%; Score 24; DB 6; Length 3541;
Best Local Similarity 64.3%; Pred. No. 17; Mismatches 20; Indels 0; Gaps 0;
Matches 36; Conservative 0;

Qy 11 TCCGTATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCAATCAATAAAGTG 66
Db 2645 TCCAGTTGCAGCCCTATTGTCTCAGATGGTCTCTAAAACCTTGAATTAATTCAGTG 2700

RESULT 12

US-10-750-185-64427
; Sequence 64427, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 64427
; LENGTH: 1049
; TYPE: DNA
; ORGANISM: Bovine 19866880882500
US-10-750-185-64427

Query Match 27.0%; Score 23.8; DB 6; Length 1049;
Best Local Similarity 57.3%; Pred. No. 13; Mismatches 32; Indels 0; Gaps 0;
Matches 43; Conservative 0;

Qy 14 TGATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCAATCAATAAAGTGCACGTG 73
Db 805 TGAATTTTATCCTAAATTTTAAAGATGTTTAAAGATGACATACATTAATGCTCATATGT 864

Qy 74 TTCTTGGTACAAAA 88
Db 865 GTTAGTTACAAA 879

RESULT 13
US-10-750-185-25851/c
; Sequence 25851, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 25851
; LENGTH: 2248
; TYPE: DNA
; ORGANISM: Bovine 19866880712906
US-10-750-185-25851

Query Match 27.0%; Score 23.8; DB 6; Length 2248;
Best Local Similarity 59.7%; Pred. No. 17; Mismatches 27; Indels 0; Gaps 0;
Matches 40; Conservative 0;

Qy 14 TGATCAGCCCTTTATCCTCACACGCTTTTCTACAATGGCAATCAATAAAGTGCACGTG 73
Db 1517 TATTTGATCCTTTTATTTTCAATGTTATATATATCTTGGTCTATAAAAAGACTGTTT 1458

Qy 74 TTCTTGG 80

Db 1457 TTCTGG 1451
|||||||

RESULT 14
US-10-793-626-419
; Sequence 419, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793.626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 419
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-419

Query Match 27.0%; Score 23.8; DB 6; Length 2301;
Best Local Similarity 72.1%; Pred. No. 17;
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 31 CTCACACGCTTTTCTACAAATGGCATTCAATAAAGTGCACGTGT 73
|||||
Db 1817 CTCAAACATTGATACAAATGGCATTCAAAAAGGGGAACGTGT 1859
|||||

RESULT 15
US-10-793-626-4069/c
; Sequence 4069, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793.626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4069
; LENGTH: 2744
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-4069

Query Match 27.0%; Score 23.8; DB 6; Length 2744;
Best Local Similarity 72.1%; Pred. No. 18;
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 31 CTCACACGCTTTTCTACAAATGGCATTCAATAAAGTGCACGTGT 73
|||||
Db 715 CTCAAACATTGATACAAATGGCATTCAAAAAGGGGAACGTGT 673
|||||

Search completed: December 6, 2005, 23:16:51
Job time : 60.0707 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 1004.45 Seconds
(without alignments)
4099.008 Million cell updates/sec

Title: US-09-980-046B-5
Perfect score: 88
Sequence: 1 cggagtcgagtcctgtatca.....cgtgttcttctgtacaaaaa 88

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_hc:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_est7:*

9: gb_ges1:*

10: gb_ges2:*

11: gb_ges3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84.8	95.4	661	AV702819	AV702819
2	83.8	95.2	639	CF129716	CF129716 UI-HF-ES0
3	82.2	93.4	640	BF685813	BF685813 602140496
4	81.6	92.7	593	BI517742	BI517742 603042174
5	81.6	92.7	594	BO689205	BO689205 AGENCOURT
6	81.6	92.7	632	BI599867	BI599867 603246815
7	81.6	92.7	652	BF339806	BF339806 602038812
8	81.6	92.7	914	BI762444	BI762444 603048836
9	80.6	91.6	661	BO067642	BO067642 AGENCOURT
10	80	90.9	236	CA431735	CA431735 UI-H-DF0-
11	80	90.9	243	AI547782	AI547782 UI-R-C3-s
12	80	90.9	251	CH310115	CH310115 170060000
13	80	90.9	271	AA542950	AA542950 ni91907.s
14	80	90.9	280	CH483092	CH483092 hw28c11.y
15	80	90.9	313	BI060666	BI060666 PM0-UT010
16	80	90.9	316	BM924345	BM924345 AGENCOURT
17	80	90.9	319	AW516688	AW516688 xDP71a03.x
18	80	90.9	330	AW950085	AW950085 EST162275
19	80	90.9	341	AI754997	AI754997 cr34a04.x
20	80	90.9	357	AA573395	AA573395 nm53b11.s
21	80	90.9	358	AA657893	AA657893 nu07h03.s
22	80	90.9	358	BM543768	BM543768 AGENCOURT

c	23	80	90.9	358	8	N29647	N29647 yw73b02.s1
	24	80	90.9	362	5	CF129432	CF129432 UI-HF-ES0
	25	80	90.9	363	5	EX118214	EX118214 BX118214
	26	80	90.9	387	6	CF272419	CF272419 AGENCOURT
	27	80	90.9	393	2	BI257614	BI257614 602967192
	28	80	90.9	393	5	BU153040	BU153040 AGENCOURT
c	29	80	90.9	399	2	EG059171	EG059171 nah46R05.
	30	80	90.9	404	1	AW023437	AW023437 df54h03.y
c	31	80	90.9	405	1	AW472903	AW472903 x822a08.x
	32	80	90.9	406	3	BM850332	BM850332 K-EST0130
	33	80	90.9	410	1	AA658408	AA658408 nu18d02.s
c	34	80	90.9	412	1	AI862050	AI862050 tw71b08.x
	35	80	90.9	425	1	AA534111	AA534111 nj73g01.s
c	36	80	90.9	428	8	T55566	T55566 yb39c05.s1
	37	80	90.9	432	1	AA535668	AA535668 nj77g12.s
c	38	80	90.9	437	3	BM672821	BM672821 UI-E-CQ0-
	39	80	90.9	439	1	AL036296	AL036296 DKF2p564N
	40	80	90.9	442	3	BM741599	BM741599 K-EST0014
c	41	80	90.9	446	5	BU728111	BU728111 UI-E-CQ0-
	42	80	90.9	449	1	AA535290	AA535290 nj79b05.s
c	43	80	90.9	450	8	N93952	N93952 zb74604.s1
	44	80	90.9	458	3	BQ029595	BQ029595 UI-H-DF0-
c	45	80	90.9	458	5	BU737184	BU737184 UI-E-DX1-

ALIGNMENTS

RESULT 1
AV702819
LOCUS
DEFINITION AV702819 ADB Homo sapiens cDNA clone ADBDGH07 5', mRNA linear EST 08-OCT-2000
ACCESSION AV702819
VERSION AV702819.1 GI:10719149
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 661)
AUTHORS Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G.,
Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Geng, Z.,
Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S.,
Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z. and Han, Z.
TITLE Homo sapiens cDNA ADB clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzgchgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Location/Qualifiers
1..661
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="ADB07"
/tissue type="Adrenal gland"
/dev stage="Adult"
/lab_host="SOLR"
/clone_lib="ADB"
/note="Vector: pbluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Query Match 96.4%; Score 84.8; DB 1; Length 661;
Best Local Similarity 97.7%; Pred. No. 7.5e-18;
Matches 86; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 CGAGTCCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAAATGGCATTCAAT 60
    |||||
Db 477 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAAATGGCATTCAAT 536
    |||||

QY 61 AAGTGCACGCTGTTCTTCTGGTACAAAAA 88
    |||||
Db 537 AAGTGCACGCTGTTCTTCTGGTAAAAAAA 564
    |||||

RESULT 2
CFI29716 639 bp mRNA linear EST 05-AUG-2003
LOCUS UI-HF-ESO-avv-p-11-0-UI.r1 NIH_MGC_213 Homo sapiens cDNA clone
DEFINITION IMAGE:30565138 5', mRNA sequence.
ACCESSION CFI29716
VERSION CFI29716.1 GI:33210245
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 639)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
JOURNAL discovery
PUBMED Genome Res. 6 (9), 791-806 (1996)
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Mary Hendrix
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: pYX-5.

FEATURES
    source
    1..639
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:30565138"
    /tissue_type="Chondrosarcoma Lung Metastasis cell lines"
    /lab_host="DH10B (T1 phage resistant)"
    /clone_lib="NIH_MGC_213"
    /note="Organ: Lung; Vector: pYX-Asc; Site 1: EcoR I;
    Site 2: Not I; The library was constructed according
    Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
    1996. Denatured RNA was size fractionated on a 1% agarose
    gel. First strand cDNA synthesis was primed with oligo-dT
    primer containing a Not I site. Double strand cDNA was
    size selected according to mRNA size fraction, ligated
    with EcoR I adaptor, digested with Not I and then cloned
    directionally into pYX-Asc vector. The library tag
    sequence located between the Not I site and the polyA tail
    is GATTAAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN
Query Match 95.2%; Score 83.8; DB 6; Length 639;
Best Local Similarity 96.6%; Pred. No. 1.6e-17;
Matches 85; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGAGTCCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAAATGGCATTCAAT 60
    |||||
Db 538 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAAATGGCATTCAAT 597
    |||||

QY 61 AAGTGCACGCTGTTCTTCTGGTACAAAAA 88
    |||||

```

```

Db 598 AAGTGCACGCTGTTCTTCINGGTAAAAA 625
    |||||

RESULT 3
BF685813 640 bp mRNA linear EST 22-DEC-2000
LOCUS 602140496F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4301807 5',
DEFINITION mRNA sequence.
ACCESSION BF685813
VERSION BF685813.1 GI:11971221
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 640)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1162 row: 0 column: 24
High quality sequence stop: 638.
Location/Qualifiers
    1..640
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:4301807"
    /tissue_type="leiomyosarcoma cell line"
    /lab_host="DH10B (phage-resistant)"
    /clone_lib="NIH_MGC_46"
    /note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:
    EcoRI; cDNA made by oligo-dT priming. Directionally cloned
    into EcoRI/XhoI sites using the following 5' adaptor:
    GGCACGAG(G). Size-selected >500bp for average insert size
    1.8kb. Library constructed by Ling Hong in the laboratory
    of Gerald M. Rubin (University of California, Berkeley)
    using ZAP-cDNA synthesis kit (Stratagene) and Superscript
    II RT (Life Technologies). Note: this is a NIH_MGC
    Library."

ORIGIN
Query Match 93.4%; Score 82.2; DB 2; Length 640;
Best Local Similarity 95.5%; Pred. No. 5.7e-17;
Matches 84; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGAGTCCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAAATGGCATTCAAT 60
    |||||
Db 535 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACAGCTTTTCTACAAATGGCATTCAAT 594
    |||||

QY 61 AAGTGCACGCTGTTCTTCTGGTACAAAAA 88
    |||||
Db 595 AAGTGCACGCTGTTCTTCTGGTACAAAAA 622
    |||||

RESULT 4
BF685813 593 bp mRNA linear EST 29-AUG-2001
LOCUS 603042174F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5182435 5',
DEFINITION mRNA sequence.
ACCESSION BF685813
VERSION BF685813.1 GI:15342534
KEYWORDS EST.

```

```

SOURCE      Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo;
1 (bases 1 to 593)
REFERENCE   NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS    National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE      Unpublished (1999)
JOURNAL
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11455 row: d column: 20
            High quality sequence stop: 561.

FEATURES   Location/Qualifiers
            1..593
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5182435"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_116"
            /note="Organ: pooled colon, kidney, stomach; Vector:
            pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
            source anonymous pool of 3 colons, age 26 yo male, 49 yo
            female, 71 yo male colon; 46 yo male kidney, and pool of 2
            stomachs, 62 yo male and 70 yo female. Library is
            oligo-dT primed and directionally cloned (EcoRV site is
            destroyed upon cloning). Average insert size 1.4 kb,
            insert size range 1-3 kb. Library is normalized and
            enriched for full-length clones and was constructed by C.
            Gruber (Invitrogen). Research Genetics tracking code
            023. Note: this is a NIH_MGC Library."

ORIGIN
Query Match      92.7%; Score 81.6; DB 3; Length 593;
Best Local Similarity 95.5%; Pred. No. 8.9e-17;
Matches 84; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGAGTCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTCTACAAATGGCATTCAAT 60
    |||||
DB 482 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTCTACAAATGGCATTCAAT 541
    |||||

QY 61 AAAGTGCACGTGTTCTTGGTACAAAAA 88
    |||||
DB 542 AAAGTGCACGTGTTCTTGGTACAAAAA 569
    |||||

RESULT 5
BQ689205      594 bp mRNA linear EST 15-JUL-2002
LOCUS         AGENCOURT_8062029 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6210960
DEFINITION    5', mRNA sequence.
ACCESSION     BQ689205
VERSION       BQ689205.1 GI:21814521
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominoidea; Homo.
REFERENCE     1 (bases 1 to 594)
AUTHORS       NIH-MGC http://mgi.nci.nih.gov/.
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov

SOURCE      Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo;
1 (bases 1 to 594)
REFERENCE   NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS    National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE      Unpublished (1999)
JOURNAL
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11455 row: m column: 03
            High quality sequence stop: 632.

FEATURES   Location/Qualifiers
            1..632
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6210960"
            /tissue_type="ductal carcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_110"
            /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
            Site 2: EcoRI; cDNA made by oligo-dT priming.
            Directionally cloned into EcoRI/XhoI sites using the
            following 5' adaptor: GGCAGCAG(G). Library constructed by
            Ling Hong in the laboratory of Gerald M. Rubin (University
            of California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies).
            Note: this is a NIH_MGC Library."

ORIGIN
Query Match      92.7%; Score 81.6; DB 5; Length 594;
Best Local Similarity 95.5%; Pred. No. 9e-17;
Matches 84; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTCTACAAATGGCATTCAAT 60
    |||||
DB 497 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCTCACACGCTTTCTACAAATGGCATTCAAT 556
    |||||

QY 61 AAAGTGCACGTGTTCTTGGTACAAAAA 88
    |||||
DB 557 AAAGTGCACGTGTTCTTGGTACAAAAA 584
    |||||

RESULT 6
B1599867      632 bp mRNA linear EST 07-SEP-2001
LOCUS         B1599867 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5294762 5',
DEFINITION    mRNA sequence.
ACCESSION     B1599867
VERSION       B1599867.1 GI:15492806
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominoidea; Homo.
REFERENCE     1 (bases 1 to 632)
AUTHORS       NIH-MGC http://mgi.nci.nih.gov/.
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
              cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
              Toshiyuki and Piero Carninci (RIKEN)
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM11745 row: m column: 03
              High quality sequence stop: 632.

FEATURES   Location/Qualifiers
            1..632
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6210960"
            /tissue_type="ductal carcinoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_110"
            /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
            Site 2: EcoRI; cDNA made by oligo-dT priming.
            Directionally cloned into EcoRI/XhoI sites using the
            following 5' adaptor: GGCAGCAG(G). Library constructed by
            Ling Hong in the laboratory of Gerald M. Rubin (University
            of California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies).
            Note: this is a NIH_MGC Library."

```


Db 626 AAGTGCACGCTGTTTCTGTCGCAAAA 653

RESULT 9
BQ067642 661 bp mRNA linear EST 02-APR-2002
LOCUS
DEFINITION AGNCOURT_6759133 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5755219
5', mRNA sequence.

ACCESSION BQ067642
VERSION BQ067642.1 GI:19896688
KEYWORDS EST.

SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM12793 row: n column: 20
High quality sequence stop: 632.

FEATURES
Location/Qualifiers
1..661
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5755219"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 91.6%; Score 80.6; DB 3; Length 661;
Best Local Similarity 94.3%; Pred. No. 2e-16;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTCATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGCGATTCAAT 60
|||||
Db 540 CGGAGCCGAGTCCTGTCATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGCGATTCAAT 599
|||||

QY 61 AAGTGCACGCTGTTTCTTGGTACAAAA 88
|||||
Db 600 AAGTGCACGCTGTTTCTGTTGTAANAAAA 627
|||||

RESULT 10
CA431735 236 bp mRNA linear EST 07-NOV-2002
LOCUS
DEFINITION UI-H-DF0-bet-n-02-0-UI.s1 NCI CGAP DFO Homo sapiens cDNA clone
UI-H-DF0-bet-n-02-0-UI 3', mRNA sequence.

ACCESSION CA431735
VERSION CA431735.1 GI:24794461
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 236)
TITLE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
Location/Qualifiers
1..236
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-DF0-bet-n-02-0-UI"
/tissue_type="Subchondral Bone"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP DFO"
/note="Organ: Bone; Vector: p7T73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP DFO is a cDNA library containing the following tissue(s): Subchondral Bone. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7T73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GTTAAAGCGTC.
TAG LIB=UI-H-DF0
TAG_SEQ=GTTAAAGCGTC"

ORIGIN
Query Match 90.9%; Score 80; DB 6; Length 236;
Best Local Similarity 94.3%; Pred. No. 2.5e-16;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGAGTCGAGTCCTGTCATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGCGATTCAAT 60
|||||
Db 99 CGGAGCCGAGTCCTGTCATCAGCCCTTTATCCTCACACGCTTTTCTACAAATGCGATTCAAT 40
|||||

QY 61 AAGTGCACGCTGTTTCTTGGTACAAAA 88
|||||
Db 39 AAGTGCACGCTGTTTCTGTTGTAANAAAA 12
|||||

RESULT 11
AI547782/c
LOCUS
DEFINITION UI-R-C3-sj-g-03-0-UI.s1 UI-R-C3 Rattus norvegicus cDNA clone
UI-R-C3-sj-g-03-0-UI 3', mRNA sequence.

ACCESSION AI547782
VERSION AI547782.1 GI:4465270
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 243)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.regen.com) This clone is also available through the I.M.A.G.E. Consortium at LLNL (info@image.llnl.gov). IMAGE ID=1768820
 Seq primer: M13 Forward
 POLYA=NO.

FEATURES
source

Location/Qualifiers
 1..243
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clones="UI-R-C3-aj-g-03-0-UI"
 /dev_stage="adult"
 /lab_hosts="DH10B (Life Technologies)"
 /clone_lib="UI-R-C3"
 /notes="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C3 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged normalized libraries from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, muscle, and 8, 12 and 18-day embryos, after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: UI-R-C3, UI-R-C2p, UI-R-C1, UI-R-C0, UI-R-A1, UI-R-E1. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within themixture. The subtracted library (UI-R-C3) was constructed as follows: PCR amplified cDNA inserts from UI-R-C2p clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C2p library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C3 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)."

ORIGIN

Query Match 90.9%; Score 80; DB 1; Length 243;
 Best Local Similarity 94.3%; Pred. No. 2.5e-16;
 Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCAGCGTTTCTACAATGGCATTCAAT 60
 Db 96 CGGAGCGGAGTCCTGTATCAGCCCTTTATCTCAGCGTTTCTACAATGGCATTCAAT 37
 Qy 61 AAGTGCACGTGTTCTTGGTACAAAA 88
 Db 36 AAGTGCACGTGTTCTTGGAAAAA 9

RESULT 12

CN310115
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

CN310115 251 bp mRNA linear EST 16-MAY-2004
 17006000604048 GRN_PRENU Homo sapiens cDNA 5', mRNA sequence.
 CN310115
 CN310115.1 GI:47326529
 EST.
 Homo sapiens (human)

REFERENCE
 AUTHORS

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS

1 (bases 1 to 251)
 Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J., Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R., Lebkowski,J and Stanton,L.W.
 Transcriptionome characterization elucidates signaling networks that control human ES cell growth and differentiation
 Nat. Biotechnol. 22 (6), 707-716 (2004)

JOURNAL
 PUBMED
 COMMENT

15146197
 Contact: Brandenberger R
 Regenerative Medicine
 Geron Corporation
 230 Constitution Drive, Menlo Park, CA 94025, USA
 Tel: 650 473 8658
 Fax: 650 473 7760
 Email: rbrandenberger@geron.com
 Insert Length: 251 Std Error: 0.00.
 Location/Qualifiers

FEATURES
source

1..251
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cell, retinoic acid and mitogen-treated hES cell line H7"
 /clone_lib="GRN_PRENU"
 /notes="Oligo dT primed, full-length enriched cDNA library from hES cell line H7 (p29) maintained in feeder-free conditions. Embryoid bodies were generated in the presence of all-trans retinoic and mitogens."

ORIGIN

Query Match 90.9%; Score 80; DB 7; Length 251;
 Best Local Similarity 94.3%; Pred. No. 2.6e-16;
 Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCAGCGTTTCTACAATGGCATTCAAT 60
 Db 162 CGGAGCGGAGTCCTGTATCAGCCCTTTATCTCAGCGTTTCTACAATGGCATTCAAT 221

Qy 61 AAGTGCACGTGTTCTTGGTACAAAA 88

Db 222 AAGTGCACGTGTTCTTGGTAAAAA 249

RESULT 13
 LOCUS
 DEFINITION

AA542950 271 bp mRNA linear EST 19-AUG-1997
 n191907.s1 NCI CGAP Pr21 Homo sapiens cDNA clone IMAGE:984252 3'-
 similar to gb:X57352 INTERFERON-INDUCIBLE PROTEIN 1-8U (HUMAN);,
 mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AA542950
 AA542950.1 GI:2291430
 EST.
 Homo sapiens (human)

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

1 (bases 1 to 271)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)


```

COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
              Emmert-Buck, M.D., Ph.D.
              cDNA Library Preparation: M. Bento Soares, Ph.D.
              cDNA Library Arrayed by: Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              www-bio.llnl.gov/bbrp/image/image.html
              Insert Length: 370 Std Error: 0.00
              Seq primer: 40ml3 fwd ET from Amersham.
FEATURES     Location/Qualifiers
             1..271
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:984252"
             /sex="male"
             /tissue type="normal prostate"
             /lab host="DH10B"
             /clone lib="NCI CGAP Pr21"
             /note="Organ: prostate; Vector: pRT3D-Pac (Pharmacia)
             with a modified polylinker; 1st strand cDNA was prepared
             from normal prostate bulk tissue, and was then primed with
             a Not I - oligo(dT) primer. Double-stranded cDNA was
             ligated to Eco RI adaptors (Pharmacia), digested with Not
             I and cloned into the Not I and Eco RI sites of the
             modified pRT3 vector. Library is not normalized. Library
             was constructed by Bento Soares and M. Fatima Bonaldo."
ORIGIN
Query Match      90.9%; Score 80; DB 1; Length 271;
Best Local Similarity 94.3%; Pred. No. 2.6e-16;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCACACGCTTTTCTACAAATGGCATTCAAT 60
    |||||
Db 88 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCACACGCTTTTCTACAAATGGCATTCAAT 29

Qy 61 AAAGTGCACGCTGTTCTTGGTACAAAA 88
    |||||
Db 28 AAAGTGCACGCTGTTCTTGGTAAAAAAA 1

RESULT 14
CN483092      280 bp mRNA linear EST 26-APR-2004
LOCUS      hw28c11.y2 Human primary human ocular pericytes. Unamplified (hw)
DEFINITION      Homo sapiens cDNA clone hw28c11 5', mRNA sequence.
ACCESSION      CN483092
VERSION      CN483092.1 GI:46564596
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominidae; Homo.
              1 (bases 1 to 280)
REFERENCE      Tsai,J.Y. and Wistow,G.
              Expressed sequence tag analysis of cultured primary human ocular
              pericytes
              Unpublished (2004)
              Contact: Wistow G
              Section on Molecular Structure and Function
              National Eye Institute
              6/331, NIH, Bethesda, MD 20892-2740, USA
              Tel: 301 402 3452
              Fax: 301 496 0078
              Email: graeme@helix.nih.gov
              Plate: 28 row: c column: 11
              Seq primer: M13RP1 reverse primer (ABI).
              Location/Qualifiers
FEATURES

```

```

1..280
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hw28c11"
/cell_type="pericytes"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone lib="Human primary human ocular pericytes.
Unamplified (hw)"
/note="Organ: Eye; Vector: pSport1; RNA was extracted from
primary human pericytes in culture. A directionally cloned
cDNA library in the pSPORT1 vector (Invitrogen) was
constructed at Bioserve Biotechnology (Laurel MD)
essentially following the protocols of the SuperScript
plasmid System full details of which are contained in the
manufacturer's instruction manual
(http://www.lifetech.com/). First strand synthesis was
carried out using a Not I primer-adaptor
[5'-pGACTAGTTCTAGATCGAGCGGCCGCTT]15-3'. cDNA was
cloned in Not I/Sal I sites. EST analysis was performed at
the NIH Intramural Sequencing Center (NISC)."
ORIGIN
Query Match      90.9%; Score 80; DB 7; Length 280;
Best Local Similarity 94.3%; Pred. No. 2.6e-16;
Matches 83; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAGTCGAGTCCTGTATCAGCCCTTTATCTCACACGCTTTTCTACAAATGGCATTCAAT 60
    |||||
Db 185 CGGAGCCGAGTCCTGTATCAGCCCTTTATCTCACACGCTTTTCTACAAATGGCATTCAAT 244

Qy 61 AAAGTGCACGCTGTTCTTGGTACAAAA 88
    |||||
Db 245 AAAGTGCACGCTGTTCTTGGTAAAAAAA 272

RESULT 15
BI060666/c      313 bp mRNA linear EST 15-JUN-2001
LOCUS      PMO-UT0103-300101-002-b09 UT0103 Homo sapiens cDNA, mRNA sequence.
DEFINITION      BI060666
ACCESSION      BI060666
VERSION      BI060666.1 GI:14468193
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominidae; Homo.
              1 (bases 1 to 313)
REFERENCE      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
              Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F.,
              Goldman,G.H., deAlvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
              Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
              O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
              Simpson,A.J.
              Shotgun sequencing of the human transcriptome with ORF expressed
              sequence tags
              Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
              10737800
              Contact: Simpson A.J.G.
              Laboratory of Cancer Genetics
              Ludwig Institute for Cancer Research
              Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
              Brazil
              Tel: +55-11-2704922
              Fax: +55-11-2707001
              Email: asimpson@ludwig.org.br
              This sequence was derived from the FAPESP/LICR Human Cancer Genome
              Project. This entry can be seen in the following URL
              (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM04t2=PM0-UT0103-
              300101-002-b09&t3=2001-01-30&t4=1)
              Seq primer: puc 13 forward

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 2856.57 Seconds
(without alignments)
6427.437 Million cell updates/sec

Title: US-09-980-046B-6
Perfect score: 323
Sequence: 1 cggggccagcctatgcct.....gtgcaogtctgtgaaaaa 323

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_in.*
- 3: gb_env.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pr.*
- 9: gb_ro.*
- 10: gb_sts.*
- 11: gb_by.*
- 12: gb_un.*
- 13: gb_vi.*
- 14: gb_btg.*
- 15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	321.4	99.5	683	CS113008	Sequence
2	321.4	99.5	683	CS113019	Sequence
3	321.4	99.5	851	BC000897	Homo sapi
4	321.4	99.5	851	BD222216	Human nuc
5	321.4	99.5	851	AX014337	Sequence
6	319.8	99.0	1421	CQ412786	Sequence
7	319.4	98.9	1095	AX281854	Sequence
8	318.4	98.6	877	CQ894717	Sequence
9	317.4	98.3	641	CQ726860	Sequence
10	317.4	98.3	21478	CQ869625	Sequence
11	317.4	98.3	143779	AC136475	Homo sapi
12	317.4	98.3	156819	AP006283	Homo sapi
13	317.4	98.3	175416	AP006286	Homo sapi
14	317.4	98.3	187160	AC138230	Homo sapi
15	315.8	97.8	647	CQ776390	Sequence
16	315.8	97.8	647	CQ833999	Sequence
17	315.8	97.8	647	CQ981463	Sequence
18	315.8	97.8	647	CS030890	Sequence

19	315.8	97.8	647	6	CS033357	Sequence
20	315.8	97.8	647	6	CS036047	Sequence
21	315.8	97.8	647	6	CS039842	Sequence
22	315.8	97.8	647	6	CS042309	Sequence
23	315.8	97.8	647	6	CS044999	Sequence
24	315.8	97.8	647	6	CS108186	Sequence
25	315.8	97.8	647	6	CS1108354	Sequence
26	315.8	97.8	647	6	CS119483	Sequence
27	315.8	97.8	647	8	HSRNA927	X84958 H. sapiens m
28	315.8	97.8	677	6	CQ768913	Sequence
29	315.8	97.8	842	6	AX397528	Sequence
30	315.8	97.8	853	6	CQ812316	Sequence
31	315.8	97.8	853	6	CQ895390	Sequence
32	315.8	97.8	853	6	CS119626	Sequence
33	315.8	97.8	853	6	AX302539	Sequence
34	315.8	97.8	853	6	AX409461	Sequence
35	315.8	97.8	853	6	HUM927A	J04164 Human inter
36	314.4	97.3	675	10	BV178634	BV178634 sqm10170
37	311.8	96.5	764	6	AR148104	Sequence
38	311.8	96.5	764	6	BD084479	BD084479 Compositi
39	311.8	96.5	764	6	AR350926	Sequence
40	311.8	96.5	764	6	AR433302	Sequence
41	311.8	96.5	764	6	AR615046	Sequence
42	311.8	96.5	764	6	AR642088	Sequence
43	311.8	96.5	764	6	AX316946	Sequence
44	302.2	93.6	427	6	CQ407796	Sequence
45	296.8	91.9	535	6	CQ923236	Sequence

ALIGNMENTS

RESULT 1 CS113008 CS113008 683 bp DNA linear PAT 24-JUN-2005
LOCUS Sequence 26 from Patent WO2005054507.
DEFINITION CS113008
ACCESSION CS113008
VERSION CS113008.1 GI:68224580

KEYWORDS :
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

REFERENCE 1

AUTHORS Corfe, B. and Chirakkal, H.

TITLE Gene screen

JOURNAL Patent: WO 2005054507-A 26 16-JUN-2005;

University of Sheffield (GB)

FEATURES

source

1..683

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match	99.5%	Score 321.4	DB 6	Length 683
Best Local Similarity	99.7%	Pred No. 3.5e-85		
Matches 322	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	1	CGGGGCCCGAGCCTATGCCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG	60	
Db	349	CGGGGCCCGAGCCTATGCCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG	408	
QY	61	CATCTCTATGACCAATGGATTACCTCTTACTGGTATTCGGCTCTGTGACAGTCTACCA	120	
Db	409	CATCTCTATGACCAATGGATTACCTCTTACTGGTATTCGGCTCTGTGACAGTCTACCA	468	
QY	121	TATTATGTTACAGATAACAGGAAACGGGGTTACTAGTAGCCGCCATAGCTCGCAA	180	
Db	469	TATTATGTTACAGATAACAGGAAACGGGGTTACTAGTAGCCGCCATAGCTCGCAA	528	
QY	181	CCTTTGCACTCCACTGTGCAATGCTGGGCCCTGCACGCTGGGGCTGTGCCCCCTGCCCT	240	

Db	529	CTTTTGACATCCACTGTGC	AAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCCCTGCCCCCT	588
Qy	241	TGCTCTCTGCCCTAGATACAGCAGTATTATATACCCACACACCTGTCTACAGTGTCAATCAAT	300	
Db	589	TGCTCTCTGCCCTAGATACAGCAGTATTATATACCCACACACCTGTCTACAGTGTCAATCAAT	648	
Qy	301	AAAGTGCACGTGCTTGTGAAAAA	323	
Db	649	AAAGTGCACGTGCTTGTGAAAAA	671	
RESULT 2				
LOCUS	CS113019	683 bp DNA	linear	PAT 24-JUN-2005
DEFINITION	Sequence 37 from Patent WO2005054507.			
ACCESSION	CS113019			
VERSION	CS113019.1	GI:68224591		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1			
AUTHORS	Corfe, B. and Chirakkal, H.			
TITLE	Gene screen			
JOURNAL	Patent: WO 2005054507-A 37 16-JUN-2005;			
University of Sheffield (GB)				
FEATURES	Location/Qualifiers			
source	1..683			
/organism="Homo sapiens"				
/mol_type="unassigned DNA"				
/db_xref="taxon:9606"				
ORIGIN				
Query Match 99.5%; Score 321.4; DB 6; Length 683;				
Best Local Similarity 99.7%; Pred. No. 3.5e-85;				
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				
Qy	1	CGGGGCCAGCGCTATGCTCCACCGCAAGTGCCTGAACATCTGGCCCTGATTCGGG	60	
Db	349	CGGGGCCAGCGCTATGCTCCACCGCAAGTGCCTGAACATCTGGCCCTGATTCGGG	408	
Qy	61	CATCCTCATGACCAATCGATTACCCCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA	120	
Db	409	CATCCTCATGACCAATCGATTACCCCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA	468	
Qy	121	TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA	180	
Db	469	TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA	528	
Qy	181	CTTTTGCACTCCACTGTGCAATGCTGGCCCTGACGCTGGGGCTGTTGCCCCCTGCCCCCT	240	
Db	529	CTTTTGCACTCCACTGTGCAATGCTGGCCCTGACGCTGGGGCTGTTGCCCCCTGCCCCCT	588	
Qy	241	TGGTCTCTGCCCTAGATACAGCAGTATTATATACCCACACACCTGTCTACAGTGTCAATCAAT	300	
Db	589	TGGTCTCTGCCCTAGATACAGCAGTATTATATACCCACACACCTGTCTACAGTGTCAATCAAT	648	
Qy	301	AAAGTGCACGTGCTTGTGAAAAA	323	
Db	649	AAAGTGCACGTGCTTGTGAAAAA	671	
RESULT 3				
LOCUS	BC000897	683 bp mRNA	linear	PRI 29-JUN-2004
DEFINITION	Homo sapiens interferon induced transmembrane protein 1 (9-27), mRNA (cDNA clone MGC:5195 IMAGE:3464598), complete cds.			
ACCESSION	BC000897			
VERSION	BC000897.1	GI:12654158		
KEYWORDS	MGC.			

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 683)
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, F., Max, S.I., Wang, J., Haieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Shevchenko, Y., Sanchez, A., Whiting, M., Madan, A., Touchman, J.W., Green, E.D., Bouffard, G.O., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krawinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED	12477932
REFERENCE	2 (bases 1 to 683)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (15-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgapsb-rc@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-ahgc.stanford.edu Contact: (Dickson, Mark) medpaxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 4 Row: b Column: 24 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4504580. Location/Qualifiers 1..683 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:5195 IMAGE:3464598" /tissue_type="Cervix, carcinoma" /clone_lib="NIH MGC_12" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" 1..683 /gene="IFITM1" /note="synonyms: 9-27, LEU13, IFI17, CD225" /db_xref="GeneID:8519" /db_xref="MIM:604456" 131..508 /gene="IFITM1" /codon_start=1 /product="interferon induced transmembrane protein 1 (9-27)"
source	
gene	
CDS	

/protein_id="AAH00897.1"
/db_xref="GeneID:8519"
/db_xref="GeneID:8519"
/db_xref="WIM:604456"
/translation="MKKEHEVAVLGAPPSTILPRSTVINIHSETSVPDHHVMSLENT
LFLNWCCLGFAPFAYSVKSRRKRVGVDVGAQYASTAKCLNIWALLIGLIMTIGFIL
LLVFGSVTVHIMLQIIQEKRGY"

ORIGIN

Query Match 99.5%; Score 321.4; DB 8; Length 683;
Best Local Similarity 99.7%; Pred. No. 3.5e-85;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 60
Db 349 CGGGGCCAGGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 408

Qy 61 CATCCTCATGACCAATGGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120
Db 409 CATCCTCATGACCAATGGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 468

Qy 121 TATTATGTTACAGATAATACAGGAAACCGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180
Db 469 TATTATGTTACAGATAATACAGGAAACCGGGTTACTAGTAGCGGCCCATAGCCTGCAA 528

Qy 181 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTCGCCCTGCCCCCT 240
Db 529 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTCGCCCTGCCCCCT 588

Qy 241 TGGTCTGCCCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCTTCAAT 300
Db 589 TGGTCTGCCCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCTTCAAT 648

Qy 301 AAGTGCACGCTGCTGTGAAAAA 323
Db 649 AAGTGCACGCTGCTGTGAAAAA 671

RESULT 4
BD222216 LOCUS 851 bp DNA linear PAT 17-JUL-2003
DEFINITION Human nucleic acid sequence originating in normal uterine tissue.
ACCESSION BD222216
VERSION BD222216.1 GI:33031986
KEYWORDS JP 2002512017-A/45.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 851)
AUTHORS Specht,T., Hinzmann,B., Schmitt,A., Pilarsky,C., Dahl,E. and Rosenthal,A.
TITLE Human nucleic acid sequence originating in normal uterine tissue
JOURNAL Patent: JP 2002512017-A 45 23-APR-2002;
METAGEN GSELSCHAFT FUER GENOM FORSCHUNG MBH
OS Homo sapiens (human)
PN JP 2002512017-A/45
PD 23-APR-2002
PF 15-APR-1999 JP 2000544691
PR 17-APR-1998 DE 198 17 946.4
PI THOMAS SPECHT, BERND HINZMANN, ARMIN SCHMITT, CHRISTIAN PILARSKY,
PI EDGAR DAHL,
PI ANDRE ROSENTHAL
PC C12N15/09, A61K38/00, A61K48/00, A61P35/00, C07K14/47, C07K16/18,
PC C12N1/19,
PC C12N1/21, C12N5/00, C12P21/08, C12Q1/68, G01N33/68, C12N15/00, PC
A61K37/02,
PC C12N5/00
CC Human nucleic acid sequence originating in normal uterine tissue
FH Key
FT source 1. .851

FT Location/Qualifiers
source 1. .851
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.5%; Score 321.4; DB 6; Length 851;
Best Local Similarity 99.7%; Pred. No. 3.6e-85;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 60
Db 500 CGGGGCCAGGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 559

Qy 61 CATCCTCATGACCAATGGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120
Db 560 CATCCTCATGACCAATGGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 619

Qy 121 TATTATGTTACAGATAATACAGGAAACCGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180
Db 620 TATTATGTTACAGATAATACAGGAAACCGGGTTACTAGTAGCGGCCCATAGCCTGCAA 679

Qy 181 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTCGCCCTGCCCCCT 240
Db 680 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTCGCCCTGCCCCCT 739

Qy 241 TGGTCTGCCCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCTTCAAT 300
Db 740 TGGTCTGCCCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCTTCAAT 799

Qy 301 AAGTGCACGCTGCTGTGAAAAA 323
Db 800 AAGTGCACGCTGCTGTGAAAAA 822

RESULT 5
AX014337 LOCUS 851 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 46 from Patent WO9954353.
ACCESSION AX014337
VERSION AX014337.1 GI:10040691
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and Pilarsky,C.
TITLE Human nucleic acid sequences of normal uterus tissue
JOURNAL Patent: WO 9954353-A 46 28-OCT-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILARSKY CHRISTIAN (DE)
FEATURES
source 1. .851
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.5%; Score 321.4; DB 6; Length 851;
Best Local Similarity 99.7%; Pred. No. 3.6e-85;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 60
Db 500 CGGGGCCAGGCCCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 559

Qy 61 CATCCTCATGACCAATGGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120

```
Db 560 CATCTCATGACCATGGATTATCTCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 619
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180
Db 620 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 679
Qy 181 CCTTTGCACTCCACTGTGCAATGTGCGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCCCT 240
Db 680 CCTTTGCACTCCACTGTGCAATGTGCGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCCCT 739
Qy 241 TGGTCTCTGCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCAATCAAT 300
Db 740 TGGTCTCTGCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCAATCAAT 799
Qy 301 AAAGTGACAGTGTGTGAAAAA 323
Db 800 AAAGTGACAGTGTGTGAAAAA 822

RESULT 6
LOCUS CQ412786 1421 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 19857 from Patent WO0170979.
ACCESSION CQ412786
VERSION CQ412786.1 GI:41320567
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1
AUTHORS Lee, J. and Lillie, J.
TITLE Genes, compositions, kits, and method for identification,
assessment, prevention, and therapy of ovarian cancer
JOURNAL Patent: WO 0170979-A 19857 27-SEP-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source Location/Qualifiers
1 .1421
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.0%; Score 319.8; DB 6; Length 1421;
Best Local Similarity 99.4%; Pred. No. 1.2e-84;
Matches 321; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGGGCCAGCCCTATGCCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db 420 CGGGGCCAGCCCTATGCCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 479
Qy 61 CATCTCATGACCATGGATTACCCCTGTACTGGTATTGGGCTCTGTGACAGTCTACCA 120
Db 480 CATCTCATGACCATGGATTACCCCTGTACTGGTATTGGGCTCTGTGACAGTCTACCA 539
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180
Db 540 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 599
Qy 181 CCTTTGCACTCCACTGTGCAATGTGCGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCCCT 240
Db 600 CCTTTGCACTCCACTGTGCAATGTGCGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCCCT 659
Qy 241 TGGTCTCTGCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCAATCAAT 300
Db 660 TGGTCTCTGCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCAATCAAT 719
Qy 301 AAAGTGACAGTGTGTGAAAAA 323
Db 720 AAAGTGACAGTGTGTGAAAAA 742
```

```
RESULT 7
LOCUS AX281854 1095 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 263 from Patent WO0177389.
ACCESSION AX281854
VERSION AX281854.1 GI:16609105
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1
AUTHORS Shiffman, D., Somogyi, R., Lawn, R., Seilhamer, J. J., Porter, G. J.,
Mikita, T. and Tai, J.
TITLE Genes expressed in foam cell differentiation
JOURNAL Patent: WO 0177389-A 263 18-OCT-2001;
Incyte Genomics, Inc. (US)
FEATURES
source Location/Qualifiers
1 .1095
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/noife="Incyte ID No: 988231.7"

ORIGIN
Query Match 98.9%; Score 319.4; DB 6; Length 1095;
Best Local Similarity 99.7%; Pred. No. 1.5e-84;
Matches 320; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CGGGGCCAGCCCTATGCCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db 775 CGGGGCCAGCCCTATGCCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 834
Qy 61 CATCTCATGACCATGGATTACCCCTGTACTGGTATTGGGCTCTGTGACAGTCTACCA 120
Db 835 CATCTCATGACCATGGATTACCCCTGTACTGGTATTGGGCTCTGTGACAGTCTACCA 894
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180
Db 895 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 954
Qy 181 CCTTTGCACTCCACTGTGCAATGTGCGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCCCT 240
Db 955 CCTTTGCACTCCACTGTGCAATGTGCGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCCCT 1014
Qy 241 TGGTCTCTGCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCAATCAAT 300
Db 1015 TGGTCTCTGCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCAATCAAT 1074
Qy 301 AAAGTGACAGTGTGTGAAA 321
Db 1075 AAAGTGACAGTGTGTGAAA 1095

RESULT 8
LOCUS CQ894717 877 bp DNA linear PAT 05-NOV-2004
DEFINITION Sequence 27 from Patent EP1471075.
ACCESSION CQ894717
VERSION CQ894717.1 GI:55467466
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1
AUTHORS Rosenthal, A. D., Pilarsky, C., Dahl, E., Specht, T., Bruemmendorf, T.,
Lichtner, R., Staub, E., Roepcke, S. and Li, X. I.
TITLE Human nucleic acid sequences expressed in pancreatic carcinomas
JOURNAL Patent: EP 1471075-A 27 27-OCT-2004;
```

```
FEATURES
  source
    Location/Qualifiers
      1. .877
      /organism="Homo sapiens"
      /mol_type="unassigned DNA"
      /db_xref="taxon:9606"

ORIGIN
Query Match      98.6%; Score 318.4; DB 6; Length 877;
Best Local Similarity 99.7%; Pred. No. 2.9e-84;
Matches 319; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACGGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 60
Db 558 CGGGGCCAGGCTATGCTCCACGGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 617
Qy 61 CATCTCATGACCAATGGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120
Db 618 CATCTCATGACCAATGGATTATCTCTGTTACTGTTATTCGGCTCTGTGACAGTCTACCA 677
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 180
Db 678 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 737
Qy 181 CCTTTGCACTCCTGCTGCAATGCTGCGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCT 240
Db 738 CCTTTGCACTCCTGCTGCAATGCTGCGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCT 797
Qy 241 TGGTCTGCTCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 798 TGGTCTGCTCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 857
Qy 301 AAAGTGCAAGTGTGTGGA 320
Db 858 AAAGTGCAAGTGTGTGGA 877

RESULT 9
LOCUS      CQ726860          641 bp      DNA      linear      PAT 03-FEB-2004.
DEFINITION Sequence 12794 from Patent WO02068579.
ACCESSION CQ726860
VERSION   CQ726860.1 GI:42291172
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
           Hominidae; Homo.
REFERENCE
  1 Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
    Kites, such as nucleic acid arrays, comprising a majority of
    humanexons or transcripts, for detecting expression and other uses
    thereof
    Patent: WO 02068579-A 12794 06-SEP-2002;
    PE Corporation (NY) (US)
  Location/Qualifiers
    1. .641
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"

ORIGIN
Query Match      98.3%; Score 317.4; DB 6; Length 641;
Best Local Similarity 99.7%; Pred. No. 5.5e-84;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACGGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 60
Db 323 CGGGGCCAGGCTATGCTCCACGGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 382
```

```
Qy 61 CAPCCTCATGACCAATGGATTACCCCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 120
Db 383 CAPCCTCATGACCAATGGATTACCCCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 442
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 180
Db 443 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 502
Qy 181 CCTTTGCACTCCTGCTGCAATGCTGCGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCT 240
Db 503 CCTTTGCACTCCTGCTGCAATGCTGCGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCT 562
Qy 241 TGGTCTGCTCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 563 TGGTCTGCTCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 622
Qy 301 AAAGTGCAAGTGTGTGGA 319
Db 623 AAAGTGCAAGTGTGTGGA 641

RESULT 10
LOCUS      CQ869625          21478 bp      DNA      linear      PAT 13-SEP-2004.
DEFINITION Sequence 46 from Patent WO2004074320.
ACCESSION CQ869625
VERSION   CQ869625.1 GI:51999484
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
           Hominidae; Homo.
REFERENCE
  1 Morris,D.W., Morris,D.W. and Malandro,M.S.
    Novel therapeutic targets in cancer
    Patent: WO 2004074320-A 46 02-SEP-2004;
    Sagres Discovery, Inc. (US)
  Location/Qualifiers
    1. .21478
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"

ORIGIN
Query Match      98.3%; Score 317.4; DB 6; Length 21478;
Best Local Similarity 99.7%; Pred. No. 8.6e-84;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACGGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 60
Db 17039 CGGGGCCAGGCTATGCTCCACGGCAAGTGCTGAAACATCTGGGCCCTGATTCGGG 17098
Qy 61 CATCCTCATGACCAATGGATTACCCCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 120
Db 17099 CATCCTCATGACCAATGGATTATCCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 17158
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 180
Db 17159 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 17218
Qy 181 CCTTTGCACTCCTGCTGCAATGCTGCGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCT 240
Db 17219 CCTTTGCACTCCTGCTGCAATGCTGCGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCT 17278
Qy 241 TGGTCTGCTCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 17279 TGGTCTGCTCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 17338
Qy 301 AAAGTGCAAGTGTGTGGA 319
Db 17339 AAAGTGCAAGTGTGTGGA 17357
```



```

/rpt_family="AT_rich"
4592. .4901
/rpt_family="AluXs"
4902. .4968
/rpt_family="AT_rich"
4979. .5091
/rpt_family="L1MB4a"
5154. .5251
/rpt_family="L1MB4a"
5620. .5683
/rpt_family="L1MB4a"
complement(5684. .5954)
/rpt_family="AluJo"
5955. .6050
/rpt_family="L1MB4a"
complement(6059. .6206)
/rpt_family="MER46C"
6349. .6578
/rpt_family="L1MB4a"
6792. .7203
/rpt_family="L2"
complement(8894. .9052)
/rpt_family="L1MC5"
9225. .9287
/rpt_family="MER68-int"
9676. .9986
/rpt_family="AluSg"
9994. .10032
/rpt_family="AT_rich"
10182. .10417
/rpt_family="L1MB5"
complement(10418. .10946)
/rpt_family="LTR59"
complement(10947. .11251)
/rpt_family="AluXs"
complement(11252. .11272)
/rpt_family="LTR59"
complement(11290. .11595)
/rpt_family="AluSp"
11702. .12014
/rpt_family="AluSg"
12047. .12137
/rpt_family="L1M4"
complement(12159. .12465)
/rpt_family="AluXs"
12466. .12605
/rpt_family="L2"
12763. .12832
/rpt_family="L2"
12898. .13715
/rpt_family="L1MB5"
13716. .14007
/rpt_family="AluSc"
14018. .14320
/rpt_family="AluXs"
14321. .14603
/rpt_family="L1MB5"

Query Match 98.3%; Score 317.4; DB 8; Length 143779;
Best Local Similarity 99.7%; Pred. No. 1.1e-83;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGCCAGGCTATGCTTCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db CGGGCCAGGCTATGCTTCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 99557

Qy 61 CATCTCTATGACCATTTGATTCACCTGTACTGTATTTCGCTCTGTGACAGTCTACCA 120
Db CATCTCTATGACCATTTGATTCATCTGTACTGTATTTCGCTCTGTGACAGTCTACCA 99558

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTTACTAGTAGCGCCCATAGCCTGCAA 180
Db TATTATGTTACAGATAATACAGGAAAAACGGGGTTTACTAGTAGCGCCCATAGCCTGCAA 99677

```

```

Qy 181 CCTTTGCACTCCACTGTGCAATGCTGCGCCCTGCACGCTGGGGCTGTGCCCCCTGCCCT 240
Db 99678 CCTTTGCACTCCACTGTGCAATGCTGCGCCCTGCACGCTGGGGCTGTGCCCCCTGCCCT 99737

Qy 241 TGGTCTCGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 99738 TGGTCTCGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 99797

Qy 301 AAGTGCACTGCTTGGA 319
Db 99798 AAGTGCACTGCTTGGA 99816

```

```

RESULT 12
AP006283/c 156819 bp DNA linear HTG 26-MAR-2003
LOCUS Homo sapiens chromosome 11 clone RP11-326C3 map 11p, WORKING DRAFT
DEFINITION SEQUENCE, 19 unordered pieces.
ACCESSION AP006283
VERSION AP006283.1 GI:29243343
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

```

```

REFERENCE
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA of 11p
JOURNAL Published Only in Database (2003)
REFERENCE
AUTHORS 2 (bases 1 to 156819)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (24-MAR-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
----- Genome Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: RP11-326C3
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 144706 bases at least Q40
Consensus quality: 150967 bases at least Q30
Consensus quality: 153657 bases at least Q20
Insert size: 155019; sum-of-contigs
Quality coverage: 6.36x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
19 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved

```

```

1 21001 contig of 21001 bp in length
21102 40714 contig of 19613 bp in length
40815 58430 contig of 17616 bp in length
66754 contig of 8224 bp in length
81134 contig of 14280 bp in length
89763 contig of 8529 bp in length
101889 contig of 12026 bp in length

```

101990 112625 contig of 10636 bp in length
112726 119802 contig of 7077 bp in length
125181 129184 contig of 5279 bp in length
125282 129184 contig of 3903 bp in length
129285 137260 contig of 7976 bp in length
137361 143542 contig of 6182 bp in length
143643 146891 contig of 3249 bp in length
146992 149568 contig of 2577 bp in length
149669 152678 contig of 3010 bp in length
152779 154135 contig of 1357 bp in length
154236 155662 contig of 1427 bp in length
155763 156819 contig of 1057 bp in length.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 21001: contig of 21001 bp in length
* 21101: Gap of 100 bp
* 21102 40714: contig of 19613 bp in length
* 40715 40814: Gap of 100 bp
* 40815 58430: contig of 17616 bp in length
* 58431 58530: Gap of 100 bp
* 58531 66754: contig of 8224 bp in length
* 66755 66854: Gap of 100 bp
* 66855 81134: contig of 14280 bp in length
* 81135 81234: Gap of 100 bp
* 81235 89763: contig of 8529 bp in length
* 89764 89863: Gap of 100 bp
* 89864 101889: contig of 12026 bp in length
* 101890 101890: Gap of 100 bp
* 101990 112625: contig of 10636 bp in length
* 112626 112725: Gap of 100 bp
* 112726 119802: contig of 7077 bp in length
* 119803 119902: Gap of 100 bp
* 125181 125181: contig of 5279 bp in length
* 125182 125281: Gap of 100 bp
* 125282 129184: contig of 3903 bp in length
* 129185 129284: Gap of 100 bp
* 129285 137260: contig of 7976 bp in length
* 137261 137360: Gap of 100 bp
* 137361 143542: contig of 6182 bp in length
* 143543 143642: Gap of 100 bp
* 143643 146891: contig of 3249 bp in length
* 146892 146991: Gap of 100 bp
* 146992 149568: contig of 2577 bp in length
* 149569 149668: Gap of 100 bp
* 149669 152678: contig of 3010 bp in length
* 152679 152778: Gap of 100 bp
* 152779 154135: contig of 1357 bp in length
* 154136 154235: Gap of 100 bp
* 154236 155662: contig of 1427 bp in length
* 155663 155762: Gap of 100 bp
* 155763 156819: contig of 1057 bp in length.

FEATURES

source
1. .156819
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11p"
/clone="RP11-326C3"
1. .21001
/note="assembly_fragment"
21102. .40714
/note="assembly_fragment"
40815. .58430
/note="assembly_fragment"
58531. .66754
/note="assembly_fragment clone_end:SP6 vector_side:left"

misc_feature 66855. .81134
/note="assembly_fragment"
misc_feature 81235. .89763
/note="assembly_fragment"
misc_feature 89864. .101889
/note="assembly_fragment"
misc_feature 101990. .112625
/note="assembly_fragment"
misc_feature 112726. .119802
/note="assembly_fragment"
misc_feature 119903. .125181
/note="assembly_fragment"
misc_feature 125282. .129184
/note="assembly_fragment clone_end:T7 vector_side:right"
misc_feature 129285. .137260
/note="assembly_fragment"
misc_feature 137361. .143542
/note="assembly_fragment"
misc_feature 143643. .146891
/note="assembly_fragment"
misc_feature 146992. .149568
/note="assembly_fragment"
misc_feature 149669. .152678
/note="assembly_fragment"
misc_feature 152779. .154135
/note="assembly_fragment"
misc_feature 154236. .155662
/note="assembly_fragment"
misc_feature 155763. .156819
/note="assembly_fragment"

ORIGIN
Query Match 98.3%; Score 317.4; DB 14; Length 156819;
Best Local Similarity 99.7%; Pred. No. 1.1e-83;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGGCCAGGCTATGCTCCACCGCCAGTGCCTGGAACATCTGGGCCCTGATTCGGG 60
Db 77394 CGGGGGCCAGGCTATGCTCCACCGCCAGTGCCTGGAACATCTGGGCCCTGATTCGGG 77335

Qy 61 CATCCTCATGACCATTCGATTCACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
Db 77334 CATCCTCATGACCATTCGATTCACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 77275

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTCCGCCCATAGCTCGCAA 180
Db 77274 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTCCGCCCATAGCTCGCAA 77215

Qy 181 CCTTGCACCTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240
Db 77214 CCTTGCACCTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 77155

Qy 241 TGGTCTCTGCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCTATTCAT 300
Db 77154 TGGTCTCTGCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCTATTCAT 77095

Qy 301 AAAGTGCACGTGCTTGGA 319
Db 77094 AAAGTGCACGTGCTTGGA 77076

RESULT 13

AP006286/c
LOCUS
DEFINITION AP006286 175416 bp DNA linear HTG 26-MAR-2003
Homo sapiens chromosome 11 clone RP13-25N22 map 11p, WORKING DRAFT
SEQUENCE, 22 unordered pieces.
ACCESSION AP006286.1 GI:29243346
VERSION AP006286.1
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

```

REFERENCE
1
AUTHORS
Hattori,M., Ishii,K., Toyoda A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE
Homo sapiens genomic DNA of 11p
JOURNAL
Published Only in Database (2003)
REFERENCE
2 (bases 1 to 175416)
AUTHORS
Hattori,M., Ishii,K., Toyoda A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE
Direct Submission
JOURNAL
Submitted (24-MAR-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Sushiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
----- Genomic Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: Humdraft11
Center clone name: RP13-25N22
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 160074 bases at least Q40
Consensus quality: 167722 bases at least Q30
Consensus quality: 171662 bases at least Q20
Insert size: 173316; sum-of-contigs
Quality coverage: 7.92x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
22 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved.
1
33131 contig of 33131 bp in length
33232 54098 contig of 20867 bp in length
54199 72037 contig of 17839 bp in length
72138 83277 contig of 11140 bp in length
83378 98599 contig of 15222 bp in length
98700 105513 contig of 6814 bp in length
105614 116229 contig of 10616 bp in length
116330 121964 contig of 5635 bp in length
122065 127627 contig of 4261 bp in length
127728 131971 contig of 4244 bp in length
132072 137506 contig of 5435 bp in length
137607 143829 contig of 6223 bp in length
143930 148539 contig of 4610 bp in length
148640 153148 contig of 4509 bp in length
153249 157509 contig of 4261 bp in length
157610 160561 contig of 2952 bp in length
160662 164778 contig of 4117 bp in length
164879 168269 contig of 3391 bp in length
168370 168369 contig of 100 bp
168370 170689 contig of 2320 bp in length
170690 170789 contig of 100 bp
170790 172833 contig of 2044 bp in length
172834 172933 contig of 100 bp
172934 174155 contig of 1222 bp in length
174156 174255 contig of 100 bp
174256 175416 contig of 1161 bp in length.

FEATURES
source
1. .175416
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11p"
/clone="RP13-25N22"
1. .33131
/note="assembly_fragment clone_end:SP6 vector_side:left"
misc_feature
33232..54098
/note="assembly_fragment"
misc_feature
54199..72037
/note="assembly_fragment"
misc_feature
72138..83277
/note="assembly_fragment"
misc_feature
83378..98599
/note="assembly_fragment"
misc_feature
98700..105513
/note="assembly_fragment clone_end:T7 vector_side:left"
misc_feature
105614..116229
/note="assembly_fragment"
misc_feature
116330..121964
/note="assembly_fragment"
misc_feature
122065..127627
/note="assembly_fragment"
misc_feature
127728..131971
/note="assembly_fragment"
misc_feature
132072..137506
/note="assembly_fragment"
misc_feature
137607..143829
/note="assembly_fragment"
misc_feature
143930..148539

```

```
/note="assembly_fragment"
148640..153148
/note="assembly_fragment"
153249..157509
/note="assembly_fragment"
157610..160561
/note="assembly_fragment"
160662..164778
/note="assembly_fragment"
164879..168269
/note="assembly_fragment"
168370..170689
/note="assembly_fragment"
170790..172833
/note="assembly_fragment"
172934..174155
/note="assembly_fragment"
174256..175416
/note="assembly_fragment"

ORIGIN

Query Match      98.3%; Score 317.4; DB 14; Length 175416;
Best Local Similarity 99.7%; Pred. No. 1.1e-83;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGCCCGCCCTATCGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATCTCGG 60
Db 75568 CGGGCCCGCCCTATCGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATCTCGG 75509

Qy 61 CATCCTCATGACCATTCGATTCACCCCTGTTACTGTGTTATCGGCTCTGTGACAGTCTACCA 120
Db 75508 CATCCTCATGACCATTCGATTCACCCCTGTTACTGTGTTATCGGCTCTGTGACAGTCTACCA 75449

Qy 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180
Db 75448 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 75389

Qy 181 CCTTTGCACCTCCTGTCATGCTGCGCCCTGACGCTGGGGCTGTGGCCCTGCCCCCT 240
Db 75388 CCTTTGCACCTCCTGTCATGCTGCGCCCTGACGCTGGGGCTGTGGCCCTGCCCCCT 75329

Qy 241 TGGTCCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 75328 TGGTCCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 75269

Qy 301 AAGTGCACGCTGCTTGGA 319
Db 75268 AAGTGCACGCTGCTTGGA 75250

RESULT 14
AC138230 AC138230 187160 bp DNA linear PRI 22-MAY-2003
LOCUS Homo sapiens chromosome 11, clone RP13-317D12, complete sequence.
DEFINITION AC138230
ACCESSION AC138230
VERSION AC138230.5 GI:30984765
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1 (bases 1 to 187160)
AUTHORS Birren,B., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
TITLE Camarata,J., Chang,J., Chazaro,K., Choepel,J., Collymore,A.,
JOURNAL Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
REFERENCE Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
AUTHORS Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
2 (bases 1 to 187160)
AUTHORS Birren,B., Nussbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,K., Choepel,J., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lander,E., Abouelleil,A., Allen,N.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Mactean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Menga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-APR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 187160)
AUTHORS Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lander,E., Abouelleil,A., Allen,N.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Mactean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Menga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-APR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 187160)
AUTHORS Birren,B., Nussbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lander,E., Abouelleil,A., Allen,N.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Mactean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Menga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (22-MAY-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 22, 2003 this sequence version replaced gi:29423936.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence submissions@genome.wi.mit.edu
----- Project Information
```

Center project name: L29013
Center clone name: 317_D12

FEATURES
source Location/Qualifiers
1.187160
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP13-317D12"
135.200
/clone_lib="RP13-13 Human Female BAC"
134.1310
/rpt_family="L2"
134.1310
/rpt_family="MIR"
complement(1311..1614)
/rpt_family="AluY"
1615.1638
/rpt_family="MIR"
1849.1938
/rpt_family="L1ME"
complement(4297..4385)
/rpt_family="L2"
4463.4561
/rpt_family="PRAM"
complement(4573..4613)
/rpt_family="MAD81"
4733.5040
/rpt_family="AluY"
complement(5094..5383)
/rpt_family="L2"
complement(5379..5505)
/rpt_family="L2"
complement(5501..5627)
/rpt_family="L2"
complement(5667..6096)
/rpt_family="L2"
6542.6666
/rpt_family="MER41B"
6699.6719
/rpt_family="GC-rich"
6917.6996
/rpt_family="C-rich"
7136.7329
/rpt_family="C-rich"
15461.15483
/rpt_family="L2"
18178.18413
/rpt_family="L2"
complement(18543..18907)
/rpt_family="MER4A"
complement(18913..19138)
/rpt_family="AluSg/x"
complement(19139..19318)
/rpt_family="MER4D"
complement(19319..19688)
/rpt_family="MER4E1"
19689.19992
/rpt_family="AluSg"
19994.20308
/rpt_family="AluSg"
complement(20309..20630)
/rpt_family="MER4E1"
20632.20763
/rpt_family="AluSg/x"
complement(20785..20882)
/rpt_family="MER4E"
complement(20886..21091)
/rpt_family="MER41E"
21092.21402
/rpt_family="AluSx"
complement(21403..21580)

repeat_region /rpt_family="MER41B"
complement(22261..22538)
repeat_region /rpt_family="MER41B"
complement(22543..22594)
repeat_region /rpt_family="MER4D"
22595..22902
repeat_region /rpt_family="AluY"
complement(22903..23195)
repeat_region /rpt_family="MER4D"
23197..23314
repeat_region /rpt_family="FLAM C"
complement(23318..23368)
repeat_region /rpt_family="MER4D"
23369..23890
repeat_region /rpt_family="MER4A"
complement(23891..24166)
repeat_region /rpt_family="MER4D"
complement(24381..24698)
repeat_region /rpt_family="MER63"
complement(25372..25668)
repeat_region /rpt_family="AluY"
26402..26493
repeat_region /rpt_family="(TG)n"
complement(27325..28044)
repeat_region /rpt_family="LTR30"
28747..28755
unsure /notes="<30 qual SNGL region"
repeat_region 29024..29047
/rpt_family="GC-rich"
repeat_region complement(29397..30103)
/rpt_family="LTR8"
repeat_region 30537..30656
/rpt_family="LTR2C"

Query Match 98.3%; Score 317.4; DB 8; Length 187160;
Best Local Similarity 99.7%; Pred. No. 1.1e-83;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGGAACATCTGGGCGCTGATTCGGG 60
Db 32833 CGGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGGAACATCTGGGCGCTGATTCGGG 32892
Qy 61 CATCCTCATGACCATTCGATTCACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120
Db 32893 CATCCTCATGACCATTCGATTCACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 32952
Qy 121 TATTATGTTACAGATAATACAGGAAACCGGGTTACTAGTACCGGCCATAGCTGCAC 180
Db 32953 TATTATGTTACAGATAATACAGGAAACCGGGTTACTAGTACCGGCCATAGCTGCAC 33012
Qy 181 CCTTTGCACCTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCT 240
Db 33013 CCTTTGCACCTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCT 33072
Qy 241 TGGTCTTCGCCCTAGATACAGCAGTTTATATCCACACACCTGTCTACAGTGTCTATTCAAT 300
Db 33073 TGGTCTTCGCCCTAGATACAGCAGTTTATATCCACACACCTGTCTACAGTGTCTATTCAAT 33132
Qy 301 AAAGTGACGCTGCTGTGA 319
Db 33133 AAAGTGACGCTGCTGTGA 33151

RESULT 15
CQ776390
LOCUS CQ776390 647 bp DNA linear PAT 11-MAR-2004
DEFINITION Sequence 76 from Patent EP1394274.
ACCESSION CQ776390
VERSION CQ776390.1 GI:45379780
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1
AUTHORS Ohtani,N., Sugita,Y., Yamaya,M., Kubo,H., Nagai,H. and Izuahara,K.
TITLE Methods of testing for bronchial asthma or chronic obstructive
JOURNAL pulmonary disease
Patent: EP 1394274-A 76 03-MAR-2004;
Genox Research, Inc. (JP)
FEATURES Location/Qualifiers
source 1..647
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 97.8%; Score 315.8; DB 6; Length 647;
Best Local Similarity 99.4%; Pred. No. 1.7e-83;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCCTGGG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
329 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCCTGGG 388
Qy 61 CATCTCATGACCAATTGGATTACCCCTGTTACTGGTATTCTGGCTCTGTGACAGTCTACCA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
389 CATCTCATGACCAATTGGATTACCCCTGTTACTGGTATTCTGGCTCTGTGACAGTCTACCA 448
Qy 121 TATTATGTTACAGATAATACAGGAAACCGGGTTACTAGTACCGGCCCATAGCCTGCAA 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
449 TATTATGTTACAGATAATACAGGAAACCGGGTTACTAGTACCGGCCCATAGCCTGCAA 508
Qy 181 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
509 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 568
Qy 241 TGGTCCTGCCCCCTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGTCAATCAAT 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
569 TGGTCCTGCCCCCTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGTCAATCAAT 628
Qy 301 AAAGTGACGTGCTTGTA 319
Db ||||||||||||||||||||
629 AAAGTGACGTGCTTGTA 647

Search completed: December 7, 2005, 04:22:04
Job time : 2859.57 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 507.782 Seconds
(without alignments)
4239.411 Million cell updates/sec

Title: US-09-980-046B-6
Perfect score: 323
Sequence: 1 cggggccaggcctatgcct.....gtgcagctgtgtgaaaaa 323

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002as.*
- 7: geneseqn2002bs.*
- 8: geneseqn2003as.*
- 9: geneseqn2003bs.*
- 10: geneseqn2003cs.*
- 11: geneseqn2003ds.*
- 12: geneseqn2004as.*
- 13: geneseqn2004bs.*
- 14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	323	100.0	323	4 AAC89737	AAC89737 Human gas
2	321.4	99.5	683	14 AEA36105	Aea36105 Human nuc
3	321.4	99.5	683	14 AEA36094	Aea36094 Human nuc
4	321.4	99.5	715	13 ADT50813	Adt50813 Cancer re
5	321.4	99.5	851	2 AA241369	Aa241369 Human nor
6	321.4	99.5	1377	13 ADT50814	Adt50814 Cancer re
7	319.8	99.0	1421	5 ADL45967	Adl45967 Human ova
8	319.4	98.9	1095	6 AAS95008	Aas95008 Human DNA
9	318.4	98.6	877	13 ADX97479	Adx97479 Pancreati
10	318.2	98.5	593	6 ABQ60272	Abq60272 Human col
11	317.4	98.3	1617	10 ADE25685	Ade25685 Human cdn
12	317.4	98.3	21478	13 ABD32562	Abd32562 Human can
13	315.8	97.8	647	6 ABQ60782	Abq60782 Human 9-2
14	315.8	97.8	647	8 ABX10351	Abx10351 DNA encod
15	315.8	97.8	647	10 ADD18965	Add18965 Human dis
16	315.8	97.8	647	12 ADJ74824	Adj74824 Marker ge
17	315.8	97.8	647	12 ADL82898	Adl82898 Human PRO
18	315.8	97.8	647	12 ADQ29609	Adq29609 Human col
19	315.8	97.8	647	13 ADR25073	Adr25073 Breast ca

20	315.8	97.8	647	13	ADR14228	Adr14228	HUMAN NF-
21	315.8	97.8	647	13	ADP24680	Adp24680	PRO polyP
22	315.8	97.8	647	13	ADU06094	Novel bro	
23	315.8	97.8	647	14	ADY14590	DNA encod	
24	315.8	97.8	647	14	ADY19747	DNA encod	
25	315.8	97.8	647	14	ADY17057	DNA encod	
26	315.8	97.8	647	14	AEA04400	Human CDN	
27	315.8	97.8	647	14	AEA23652	Human PRO	
28	315.8	97.8	647	14	AEA23820	Human PRO	
29	315.8	97.8	677	12	ADI26257	Human Fra	
30	315.8	97.8	842	6	ABK46192	cDNA enco	
31	315.8	97.8	853	5	ABA83110	IFN-induc	
32	315.8	97.8	853	6	ABK83770	Human CDN	
33	315.8	97.8	853	6	ABK64497	Human ben	
34	315.8	97.8	853	6	ABN95610	Gene #210	
35	315.8	97.8	853	10	ADH28927	Human chr	
36	315.8	97.8	853	10	ADK60882	Ovarian c	
37	315.8	97.8	853	13	ADS64446	Human int	
38	315.8	97.8	853	14	ACL54036	Human col	
c	39	315.2	97.6	527	6	ABQ60421	Human col
c	40	311.8	96.5	764	2	AAV69033	DNA molec
c	41	311.8	96.5	764	3	AAC80995	Human bre
c	42	311.8	96.5	764	6	AAS99841	Breast tu
c	43	311.8	96.5	764	6	ABK46885	Human bre
c	44	311.8	96.5	764	8	ADA11362	Human bre
c	45	311.8	96.5	764	10	ADC15335	Human bre

ALIGNMENTS

RESULT 1
AAC89737
ID AAC89737 standard; cdna; 323 BP.
AC AAC89737;
XX 12-MAR-2001 (first entry)
XX Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 6.
DE Human; cytostatic; immunomodulator; immunostimulant; vulnery;
KW anti-inflammatory; neuroprotective; antibacterial; gene therapy;
KW gastrointestinal inflammation; immune system disorder; genetic disorder;
KW cancer; autoimmune disorder; infection; wound healing; ss.
XX Homo sapiens.
XX WO200073324-A2.
XX 07-DEC-2000.
XX 01-JUN-2000; 2000WO-US015191.
XX 01-JUN-1999; 99US-0137058P.
XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.
PI Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Haseel KW;
XX WPI; 2001-061508/07.
PT New polynucleotides and polypeptides, useful in gene therapy and in
PT diagnosing a pathological condition, e.g. for modulating gene expression
PT in gastrointestinal inflammation, or for treating cancers or genetic
XX disorders.
PS Claim 1; Page 85; 108pp; English.
XX The present sequence is one of a number of isolated human polynucleotides
CC which are useful in gene therapy, and for diagnosing a pathological
CC condition or a susceptibility to it. In particular, the polynucleotides
CC are useful for modulating gene expression in gastrointestinal

Db 469 TATTATGTTACAGATAATACAGGAAAAACGGGTTACTAGTAGCGGCCATAGCCTGCAA 528
 Qy 181 CCTTGGCACTCCACTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCT 240
 Db 529 CCTTGGCACTCCACTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCT 588
 Qy 241 TGGTCTGCTCCCTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGTCAATCAAT 300
 Db 589 TGGTCTGCTCCCTAGATACAGCAGTTTATACCCACACACACCTGTCTACAGTGTCAATCAAT 648
 Qy 301 AAAGTGCACTGCTGTGAAAAA 323
 Db 649 AAAGTGCACTGCTGTGAAAAA 671

RESULT 3

AEA36094

ID AEA36094 standard; DNA; 683 BP.

XX AC AEA36094;

XX DT 25-AUG-2005 (first entry)

XX XX

XX Human nucleic acid sequence #26.

XX DE

XX Screening; gene expression; colorectal tumor; colitis; Crohns disease;

XX KW irritable bowel syndrome; gastrointestinal disease; cytostatic;

XX KW gastrointestinal-gen.; antiinflammatory; ds.

XX XX

XX OS Homo sapiens.

XX XX

XX FN WO2005054507-A2.

XX XX

XX PD 16-JUN-2005.

XX XX

XX PF 03-DEC-2004; 2004WO-GB005078.

XX XX

XX PR 04-DEC-2003; 2003GB-00028048.

XX XX

XX PA (UYSH-) UNIV SHEFFIELD.

XX XX

XX PI Corfe B, Chirakkal H;

XX XX

XX DR WPI; 2005-435407/44.

XX XX

XX Screening for nucleic acid molecules exhibiting altered expression in

XX PT cells grown in the presence of butyrate, and detection of the nucleic

XX PT acid molecules or the encoded polypeptides in diagnosing colorectal

XX PT cancer.

XX XX

XX PS Disclosure; Page 55; 266pp; English.

XX XX

XX The invention relates to a method of screening for nucleic acid molecules

XX CC that show altered expression in a first cell sample comprising comparing

XX CC the gene expression profile of the sample with that of a second reference

XX CC sample, where the first sample has been grown in the presence of butyrate

XX CC or a related carbon source from which butyrate is directly or indirectly

XX CC derived, but the reference sample has not. The invention also relates to

XX CC a method of detecting at least one nucleic acid molecule associated with

XX CC the initiation and/or progression of colorectal cancer in an animal,

XX CC comprising providing a biological sample comprising at least one cell to

XX CC be tested, contacting the sample with a ligand (preferably a hybridizing

XX CC nucleic acid molecule) which binds to at least one nucleic acid and

XX CC detecting the presence of at least one molecule in the sample, a method

XX CC of detecting at least one polypeptide associated with the initiation

XX CC and/or progression of colorectal cancer in an animal comprising providing

XX CC a biological sample comprising at least one cell to be tested, contacting

XX CC the sample with at least one ligand that specifically binds at least one

XX CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising

XX CC an amino acid sequence which varies by the addition, deletion or

XX CC substitution of at least one amino acid residue and detecting the

XX CC presence of the polypeptide in the sample, a method of screening for

XX CC

CC agents that modulate the activity of at least one polypeptide encoded by
 CC a gene associated with the initiation and/or progression of colorectal
 CC cancer comprising forming a preparation comprising at least one
 CC polypeptide encoded by a nucleic acid or a variant polypeptide comprising
 CC at least one addition, deletion or substitution and at least one agent to
 CC be tested and determining the activity of the agent with respect to
 CC activity of the polypeptide, and an antibody or its effective binding
 CC portion identified by the method, for use as a pharmaceutical. The
 CC methods are useful for screening for nucleic acid molecules that show
 CC altered expression in a cell sample, and for detecting a nucleic acid and
 CC a polypeptide respectively, that are associated with the initiation
 CC and/or progression of colorectal cancer and are useful for detecting or
 CC monitoring colorectal cancer, especially adenocarcinoma. The methods are
 CC also useful for screening for agents that modulate the activity of at
 CC least one polypeptide encoded by a gene associated with the initiation
 CC and/or progression of cancer, where agents identified by the method are
 CC useful for treating colorectal cancer. The methods could also be used to
 CC detect or monitor other conditions such as colitis, Crohn's disease or
 CC irritable bowel syndrome, as a screening tool for fiber consumption, as
 CC an assay for colon microflora functionality or for early detection of pre
 CC -cancerous growth. This sequence represents a human nucleic acid
 CC identified by the screening method of the invention.

XX SQ Sequence 683 BP; 167 A; 214 C; 155 G; 147 T; 0 U; 0 Other;

Query Match 99.5%; Score 321.4; DB 14; Length 683;

Best Local Similarity 99.7%; Pred. No. 4.9e-91;

Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATCTGGG 60

Db 349 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATCTGGG 408

Qy 61 CATCTCATGACCAATTGGATTACCCCTGTTACTGGTATTCTGGCTCTGTGACAGTCTACCA 120

Db 409 CATCTCATGACCAATTGGATTACCCCTGTTACTGGTATTCTGGCTCTGTGACAGTCTACCA 468

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180

Db 469 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 528

Qy 181 CCTTTGCACCTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCT 240

Db 529 CCTTTGCACCTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCT 588

Qy 241 TGGTCTGCTCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300

Db 589 TGGTCTGCTCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 648

Qy 301 AAAGTGCACTGCTGTGAAAAA 323

Db 649 AAAGTGCACTGCTGTGAAAAA 671

RESULT 4

ADT50813

ID ADT50813 standard; DNA; 715 BP.

XX AC ADT50813;

XX DT 13-JAN-2005 (first entry)

XX XX

XX Cancer related nucleic acid sequence #117.

XX DE

XX ds; gene; cytostatic; gene therapy; vaccine; diagnosis; breast; colon;

XX KW lung; ovarian; prostate; cancer.

XX XX

XX OS Homo sapiens.

XX XX

XX WO2004092338-A2.

XX XX

XX PD 28-OCT-2004.

XX XX

PF 12-APR-2004; 2004WO-US011104.
XX
XX
PR 11-APR-2003; 2003US-0462399P.
PR 01-JUL-2003; 2003US-0484333P.
XX
XX (DIAD-) DIADEXUS INC.
XX
XX Macina RA, Turner LR, Sun Y, Tam A;
XX WPI; 2004-766851/75.
DR
XX
XX New cancer specific nucleic acid (CasNA) molecules, useful for
PT diagnosing, monitoring the presence of, or treating a patient with
PT breast, colon, lung, ovarian, or prostate cancer.
XX
XX Claim 1; SEQ ID NO 117; 891pp; English.
PS
XX The invention relates to an isolated nucleic acid molecule (I)
CC selectively hybridizing to, or comprising at least 95% sequence identity
CC to, any of the 362 nucleotide sequences fully defined in the
CC specification. The nucleic acid molecules and polypeptides are useful for
CC diagnosing, monitoring the presence of, or treating a patient with
CC breast, colon, lung, ovarian, or prostate cancer. This sequence
CC corresponds to a nucleic acid of the invention.
XX
SQ Sequence 715 BP; 160 A; 220 C; 176 G; 159 T; 0 U; 0 Other;
Query Match 99.5%; Score 321.4; DB 13; Length 715;
Best Local Similarity 99.7%; Pred. No. 5e-91;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGGGGCCAGGCCCTATGCTCCACCGCCAAAGTGCTGCAACATCTGGGCCCTGATCTGGG 60
DB 393 CGGGGGCCAGGCCCTATGCTCCACCGCCAAAGTGCTGCAACATCTGGGCCCTGATCTGGG 452
QY 61 CATCTCATGACCATTTGGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120
DB 453 CATCTCATGACCATTTGGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 512
QY 121 TATTATGTTACAGATAATACAGGAAAACGGGGTTACTAGTACCGGCCCATAGCTGCAA 180
DB 513 TATTATGTTACAGATAATACAGGAAAACGGGGTTACTAGTACCGGCCCATAGCTGCAA 572
QY 181 CCTTTGCACTCCACTGTGCAATGTGGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCT 240
DB 573 CCTTTGCACTCCACTGTGCAATGTGGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCT 632
QY 241 TGGTCTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
DB 633 TGGTCTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 692
QY 301 AAGTGCACGTGCTTGTGAAAAA 323
DB 693 AAGTGCACGTGCTTGTGAAAAA 715
RESULT 5
AAZ41369
ID AAZ41369 standard; cDNA; 851 BP.
XX
XX AAZ41369;
XX
XX 19-JAN-2000 (first entry)
XX Human normal uterus tissue derived cDNA 45.
DE Human; uterus; cancer; treatment; anticancer; cytostatic; gene therapy;
KW EST; expressed sequence tag; ss.
XX Homo sapiens.
OS
XX DE19817946-A1.
XX

PD 21-OCT-1999.
XX
XX 17-APR-1998; 98DE-01017946.
XX
PR 17-APR-1998; 98DE-01017946.
XX
XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX WPI; 1999-591956/51.
DR
XX
XX New nucleic acid sequences expressed in normal uterine tissues, and
PT derived polypeptides, for treatment of uterine cancer and identification
PT of therapeutic agents.
XX
XX Claim 3; Page 108; 154pp; German.
PS
XX This invention describes novel cDNA sequences (A) highly expressed in
CC normal uterine tissue which can have anticancer and cytostatic activity
CC and can be used for gene therapy. (A) are used (i) for recombinant
CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)
CC are used (i) to identify agents suitable for treatment of uterine cancer;
CC (ii) directly for treating this form of cancer (including expression from
CC gene therapy vectors) and (iii) for generation of specific antibodies.
CC (A) are identified by assembling ESTs (expressed sequence tags) from a
CC particular tissue type before comparison of expression patterns. This
CC allows a significantly longer fragment of the gene to be revealed, so
CC should reduce the number of failures associated with the fact that ESTs
CC from different libraries may represent different parts of the same
CC unknown gene, distorting the estimated frequency of occurrence in a
CC particular tissue. AAZ41325-241385 represent the human uterine tissue
CC derived cDNA fragments of the invention which encode the protein
CC fragments represented in AAY59838-Y59892
XX
SQ Sequence 851 BP; 225 A; 248 C; 211 G; 167 T; 0 U; 0 Other;
Query Match 99.5%; Score 321.4; DB 2; Length 851;
Best Local Similarity 99.7%; Pred. No. 5.4e-91;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGGGGCCAGGCCCTATGCTCCACCGCCAAAGTGCTGCAACATCTGGGCCCTGATCTGGG 60
DB 500 CGGGGGCCAGGCCCTATGCTCCACCGCCAAAGTGCTGCAACATCTGGGCCCTGATCTGGG 559
QY 61 CATCTCATGACCATTTGGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120
DB 560 CATCTCATGACCATTTGGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 619
QY 121 TATTATGTTACAGATAATACAGGAAAACGGGGTTACTAGTACCGGCCCATAGCTGCAA 180
DB 620 TATTATGTTACAGATAATACAGGAAAACGGGGTTACTAGTACCGGCCCATAGCTGCAA 679
QY 181 CCTTTGCACTCCACTGTGCAATGTGGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCT 240
DB 680 CCTTTGCACTCCACTGTGCAATGTGGCCCTGCAAGCTGGGGCTGTTGCCCTGCCCT 739
QY 241 TGGTCTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
DB 740 TGGTCTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 799
QY 301 AAGTGCACGTGCTTGTGAAAAA 323
DB 800 AAGTGCACGTGCTTGTGAAAAA 822
RESULT 6
ADT50814
ID ADT50814 standard; DNA; 1377 BP.
XX
XX ADT50814;
XX
XX 13-JAN-2005 (first entry)
DT

```
XX DE Cancer related nucleic acid sequence #118.
XX DE ds; gene; cytostatic; gene therapy; vaccine; diagnosis; breast; colon;
XX KW lung; ovarian; prostate; cancer.
XX KW Homo sapiens.
XX OS WO2004092338-A2.
XX PN 28-OCT-2004.
XX PD
XX PF 12-APR-2004; 2004WO-US011104.
XX PR 11-APR-2003; 2003US-0462399P.
XX PR 01-JUL-2003; 2003US-0484333P.
XX PA (DIAD-) DIADEXUS INC.
XX PI Macina RA, Turner LR, Sun Y, Tam A;
XX DR WPI; 2004-766851/75.
XX DR New cancer specific nucleic acid (CaSNA) molecules, useful for
PT diagnosing, monitoring the presence of, or treating a patient with
PT breast, colon, lung, ovarian, or prostate cancer.
XX PT
XX PT Claim 1; SEQ ID NO 118; 891pp; English.
XX PS
XX CC The invention relates to an isolated nucleic acid molecule (I)
XX CC selectively hybridizing to, or comprising at least 95% sequence identity
XX CC to, any of the 362 nucleotide sequences fully defined in the
XX CC specification. The nucleic acid molecules and polypeptides are useful for
XX CC diagnosing, monitoring the presence of, or treating a patient with
XX CC breast, colon, lung, ovarian, or prostate cancer. This sequence
XX CC corresponds to a nucleic acid of the invention.
XX CC
XX CC Sequence 1377 BP; 346 A; 403 C; 356 G; 263 T; 0 U; 9 Other;
XX SQ
XX Query Match 99.5%; Score 321.4; DB 13; Length 1377;
XX Best Local Similarity 99.7%; Pred. No. 6.8e-91;
XX Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 CGGGCCCGAGGCTATGCTCTCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 60
Db 845 CGGGCCCGAGGCTATGCTCTCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 904
Oy 61 CATCTCTATGACCATTTGGATTACCCCTGTTACTGTGTTATTCGGCTCTGTGACAGTCTACCA 120
Db 905 CATCTCTATGACCATTTGGATTACCCCTGTTACTGTGTTATTCGGCTCTGTGACAGTCTACCA 964
Oy 121 TATTATGTTACAGATAATACAGGAAAAAGGGGTTACTAGTAGCGGCCCATAGCTGCAG 180
Db 965 TATTATGTTACAGATAATACAGGAAAAAGGGGTTACTAGTAGCGGCCCATAGCTGCAG 1024
Oy 181 CTTTGTGCACTCCACGTGCAATGCTGGCCCTGTGACGCTGGGCTGTGCCCTGCCCCCT 240
Db 1025 CTTTGTGCACTCCACGTGCAATGCTGGCCCTGTGACGCTGGGCTGTGCCCTGCCCCCT 1084
Oy 241 TGGTCTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 1085 TGGTCTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 1144
Oy 301 AAGTGCACGCTGTTGTGAAAAA 323
Db 1145 AAGTGCACGCTGTTGTGAAAAA 1167
XX RESULT 7
XX ADL45967
XX ID ADL45967 standard; DNA; 1421 BP.
XX AC ADL45967;
```

```
XX 20-MAY-2004 (first entry)
XX Human ovarian cancer DNA marker #19857.
XX DE Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX KW Homo sapiens.
XX OS WO200170979-A2.
XX PN 27-SEP-2001.
XX PD
XX PF 21-MAR-2001; 2001WO-US009126.
XX PR 21-MAR-2000; 2000US-0191031P.
XX PR 25-MAY-2000; 2000US-0207124P.
XX PR 15-JUN-2000; 2000US-0211940P.
XX PR 07-JUL-2000; 2000US-0216820P.
XX PR 25-JUL-2000; 2000US-0220661P.
XX PR 21-DEC-2000; 2000US-0257672P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Lee J, Lillie J;
XX DR WPI; 2001-611502/70.
XX PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
XX PT cancer cells as compared to their normal non-cancerous ovarian cells are
XX PT used to characterize stage, grade, histological type of ovarian cancer.
XX PS Disclosure; SEQ ID NO 19857; 106pp; English.
XX CC The invention relates to nucleic acid markers which are overexpressed in
XX CC ovarian cancer cells as compared to their expression in normal (i.e. non-
XX CC cancerous) ovarian cells. The invention also relates to polypeptides
XX CC encoded by the markers, antibodies that selectively bind to the
XX CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
XX CC of developing ovarian cancer involving inhibiting expression of a gene
XX CC corresponding to a marker of the invention and a method of treating a
XX CC patient afflicted with ovarian cancer comprising providing to cells of
XX CC the patient an antisense oligonucleotide complementary to a marker of the
XX CC invention. The markers are useful for assessing if a patient is afflicted
XX CC with ovarian cancer, which involves comparing the level of expression of
XX CC a marker in a patient sample and a normal level of expression of the
XX CC marker in a control non-ovarian cancer sample. A difference between the
XX CC expression levels indicates ovarian cancer. The level of expression of a
XX CC marker corresponds to a secreted protein or to a transcribed
XX CC polynucleotide or its portion. The level of expression of the marker is
XX CC assessed by detecting the presence in the sample, a protein or protein
XX CC fragment corresponding to the marker. The presence of protein or protein
XX CC fragment is detected using an antibody that specifically binds with the
XX CC protein or protein fragment. Alternatively, the level of expression of
XX CC the marker is assessed by detecting the presence of a transcribed
XX CC polynucleotide which anneals with the marker or anneals with a portion of
XX CC the polynucleotide comprising the marker, under stringent conditions. The
XX CC marker is also used for monitoring the progression of ovarian cancer in a
XX CC patient which involves detecting expression of the marker in a patient
XX CC sample at a first point in time, repeating the method at a subsequent
XX CC time and comparing the level of expression. The method is carried out
XX CC using an ovarian tissue sample. A composition comprising a marker,
XX CC polypeptide or antibody of the invention is used to treat ovarian cancer.
XX CC This sequence represents a human ovarian cancer DNA marker of the
XX CC invention. Note: The sequence data for this patent did not form part of
XX CC the printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 1421 BP; 364 A; 391 C; 328 G; 338 T; 0 U; 0 Other;
```

Query Match 99.0%; Score 319.8; DB 5; Length 1421;
Best Local Similarity 99.4%; Pred. No. 2.2e-90;
Matches 321; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	CGGGCCAGCCCTATGCTCCACGGCCNAGTGCCTGAACATCTGGGCCCTGATTCCTGGG	60
Db	420	CGGGCCAGCCCTATGCTCCACGGCCAAAGTGCTGAACATCTGGGCCCTGATTCCTGGG	479
QY	61	CATCCTCATGACCAATGGATTACCCCTGTTACTTGGTATTTCGGCTCTGTGACAGTCTACCA	120
Db	480	CATCCTCATGACCAATGGATTATCCTTGTTACTTGGTATTTCGGCTCTGTGACAGTCTACCA	539
QY	121	TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA	180
Db	540	TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA	599
QY	181	CTTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGGCCCTGCCCCCT	240
Db	600	CTTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGGCCCTGCCCCCT	659
QY	241	TGGTCTGCCCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCAATCAAT	300
Db	660	TGGTCTGCCCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCAATCAAT	719
QY	301	AAAGTGACGCTGCTTGTAAGAAA	323
Db	720	AAAGTGACGCTGCTTGTAAGAAA	742
RESULT 8			
ID	AAS95008 standard; DNA; 1095 BP.		
XX	AC	AAS95008;	
XX	DT	14-FEB-2002 (first entry)	
XX	DE	Human DNA sequence #263 expressed during foam cell differentiation.	
XX	KW	Human; foam cell differentiation; atherosclerosis; cerebral stroke;	
XX	KW	cardiovascular disorder; coronary artery disease; gene therapy; ds.	
XX	OS	Homo sapiens.	
XX	PN	WO200177389-A2.	
XX	PD	18-OCT-2001.	
XX	PF	04-APR-2001; 2001WO-US011128.	
XX	PR	05-APR-2000; 2000US-0195106P.	
XX	PA	(INCY-) INCYTE GENOMICS INC.	
XX	PI	Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;	
XX	PI	Tai J;	
XX	DR	WPI; 2002-010925/01.	
XX	PT	Composition useful for diagnosis of conditions, disorders or diseases	
XX	PT	associated with atherosclerosis, comprises several polynucleotides that	
XX	PT	are differentially expressed in foam cell development.	
XX	PS	Claim 1; Page 303; 315pp; English.	
XX	CC	The present invention relates to the isolation of human polynucleotide	
XX	CC	sequences that are differentially expressed during foam cell	
XX	CC	differentiation. The polynucleotide sequences of the invention or a	
XX	CC	composition comprising these polynucleotides are useful as a high	
XX	CC	throughput method for detecting altered expression of one or more	
XX	CC	polynucleotides in a sample. The polynucleotides can be used in the	
XX	CC	diagnosis of disorders associated with foam cell development such as	
XX	CC	atherosclerosis, cerebral stroke, and cardiovascular disorders such as	
XX	CC	coronary artery disease. The polynucleotide sequences can also be used as	
XX	CC	PCR primers and probes. The polynucleotides of the invention are also	
XX	CC	useful in gene therapy. AAS94746-AAS95021 represent the human	
CC	polynucleotide sequences of the invention which are differentially		
CC	expressed during foam cell differentiation		
XX	XX	Sequence 1095 BP; 267 A; 326 C; 268 G; 234 T; 0 U; 0 Other;	
QY	Query Match	98.9%; Score 319.4; DB 6; Length 1095;	
Db	Best Local Similarity	99.7%; Pred. No. 2.6e-90;	
XX	Matches 320; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
QY	1	CGGGCCAGCCCTATGCTCCACCGCCAAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG	60
Db	775	CGGGCCAGCCCTATGCTCCACCGCCAAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG	834
QY	61	CATCCTCATGACCAATGGATTACCCCTGTTACTTGGTATTTCGGCTCTGTGACAGTCTACCA	120
Db	835	CATCCTCATGACCAATGGATTATCCTTGTTACTTGGTATTTCGGCTCTGTGACAGTCTACCA	894
QY	121	TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA	180
Db	895	TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA	954
QY	181	CTTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGGCCCTGCCCCCT	240
Db	955	CTTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGGCCCTGCCCCCT	1014
QY	241	TGGTCTGCCCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCAATCAAT	300
Db	1015	TGGTCTGCCCCCTAGATACAGCAGTTTATATCCACACACACCTGTCTACAGTGTCAATCAAT	1074
QY	301	AAAGTGACGCTGCTTGTAAGAAA	321
Db	1075	AAAGTGACGCTGCTTGTAAGAAA	1095
RESULT 9			
ID	ADX97479 standard; DNA; 877 BP.		
XX	AC	ADX97479;	
XX	DT	21-APR-2005 (first entry)	
XX	DE	Pancreatic cancer associated human protein encoding DNA, SEQ ID 27.	
XX	KW	pancreas tumor; cytostatic; gene; ds.	
XX	OS	Homo sapiens.	
XX	PN	EPI471075-A2.	
XX	PD	27-OCT-2004.	
XX	PF	31-MAR-2004; 2004EP-00090124.	
XX	PR	31-MAR-2003; 2003DE-01015834.	
XX	PA	(HINZ/) HINZMANN B.	
XX	PA	(ROSE/) ROSENTHAL A.	
XX	PA	(PILA/) PILARSKY C.	
XX	PA	(DAHL/) DAHL E.	
XX	PA	(SPEC/) SPECHT T.	
XX	PA	(LICH/) LICHTNER R.	
XX	PI	Rosenthal A, Pilarisky C, Dahl E, Specht T, Bruemendorf T;	
XX	PI	Lichtner R, Staub E, Roepcke S, Li X,	
XX	DR	WPI; 2004-768082/76.	
XX	DR	P-PSDB; ADX97550.	
XX	PT	New nucleic acid differentially expressed in pancreatic tumor tissue, for	
XX	PT	use as diagnostic agents and in screening for therapeutic agents.	
XX	PS	Claim 1; SEQ ID NO 27; 28pp; German.	

XX The invention relates to a novel human nucleic acid sequence of the
 CC pancreas and its encoded protein. The invention further comprises:
 CC proteins and peptides, preferably isolated, that contain a sequence
 CC encoded by the novel nucleic acid; and methods for diagnosis and
 CC treatment of pancreatic cancer, using a substance that inhibits or binds
 CC to the protein or its DNA, including: an antisense oligonucleotide, short
 CC interfering RNA or ribozyme directed against the pancreatic protein, an
 CC organic molecule, particularly having a molecular weight below 5000,
 CC especially 300, that binds to the pancreatic DNA, an aptamer or
 CC (monoclonal) antibody, preferably human or humanized, that binds to the
 CC pancreatic DNA, or an anti-idiotypic antibody raised against the
 CC monoclonal antibody, any of which may be derivatized with a reporter
 CC group, cytotoxic compound, immunostimulant and/or radioisotope. The novel
 CC human pancreatic proteins and their encoding DNA have cytostatic
 CC activity. The novel sequences are useful for inhibiting transcription
 CC and/or expression of genes and proteins associated with pancreatic
 CC cancer. This polynucleotide sequence represents the DNA encoding one of
 CC the novel human pancreatic proteins of the invention. Note: This sequence
 CC is not shown in the specification, it has been electronically downloaded
 CC from a DVD-ROM provided with this specification by the European Patent
 CC Office.

XX Sequence 877 BP; 213 A; 264 C; 219 G; 181 T; 0 U; 0 Other;

Query Match 98.6%; Score 318.4; DB 13; Length 877;
 Best Local Similarity 99.7%; Pred. No. 4.9e-90;
 Matches 319; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACGCCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
 Db 558 CGGGGCCAGGCTATGCTCCACGCCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 617
 Qy 61 CATCTCATGACCATGTGATTCACCTGTACTGTGTTACTCGGCTCTGTGACAGTCTACCA 120
 Db 618 CATCTCATGACCATGTGATTCACCTGTACTGTGTTACTCGGCTCTGTGACAGTCTACCA 677
 Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAG 180
 Db 678 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAG 737
 Qy 181 CTTTGTGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTGCCCCCTGCCCT 240
 Db 738 CTTTGTGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTGCCCCCTGCCCT 797
 Qy 241 TGGTCTGCCCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCTCAAT 300
 Db 798 TGGTCTGCCCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCTCAAT 857
 Qy 301 AAAGTGCACTGCTTGTGCA 320
 Db 858 AAAGTGCACTGCTTGTGCA 877

RESULT 10

ABQ60272/c

ID ABQ60272 standard; cDNA; 593 BP.

XX AC ABQ60272;

XX 02-AUG-2002 (first entry)

DE Human colon cancer related nucleotide sequence SEQ ID NO:3967.

XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;

KW genetic analysis; diagnostic; antisense therapy; gene; ss.

XX Homo sapiens.

XX WO200229086-A2.

XX 11-APR-2002.

XX

PF 02-OCT-2001; 2001WO-US030732.

XX 02-OCT-2000; 2000US-0237271P.

XX (FARB) BAYER CORP.

PA Burgess C, Astle JH, Carroll B, Catino TJ, Dwivedi P, Molino GA;

PI Thiagalingam A, Lewis ME;

XX WPI; 2002-426115/45.

XX New isolated nucleic acid that is differentially expressed in cancer

PT tissues useful for determining the presence of colon cancer in a cell or

PT tissue type, and in antisense therapy.

XX Claim 1; Fig 1; 796pp; English.

XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially

CC expressed in cancer tissues. ABQ56306 to ABQ59004 represent proteins

CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be

CC used in antisense therapy. An antibody immunoreactive with a polypeptide

CC encoded by (I) is useful for detecting cancer in a patient sample, and

CC for detecting the presence or absence of a polynucleotide encoded by a

CC nucleic acid which hybridizes to (I) in a cell. A probe/primer derived

CC from (I) can be used for determining the presence of a nucleic acid which

CC hybridizes to (I), and for determining the phenotype of cells in a sample

CC of cells from a patient. (I) is useful for determining the presence of

CC colon cancer in a cell or tissue type, for determining the presence or

CC state of other type of cancer, in antisense therapy, to generate

CC macroarrays on a solid surface, to identify a chromosome on which the

CC corresponding gene resides, and in tissue profiling, forensics, genetic

CC analysis, mapping and diagnostic applications. (I) can be used to raise

CC antibodies, and to screen for peptide analogues and antagonists

XX Sequence 593 BP; 143 A; 122 C; 164 G; 143 T; 0 U; 21 Other;

SQ Query Match 98.5%; Score 318.2; DB 6; Length 593;

Best Local Similarity 99.1%; Pred. No. 4.8e-90;

Matches 320; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTCCACGCCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60

Db 342 CGGGGCCAGGCTATGCTCCACGCCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 283

Qy 61 CATCTCATGACCATGTGATTCACCTGTACTGTGTTACTCGGCTCTGTGACAGTCTACCA 120

Db 282 CATCTCATGACCATGTGATTCACCTGTACTGTGTTACTCGGCTCTGTGACAGTCTACCA 223

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAG 180

Db 222 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAG 163

Qy 181 CTTTGTGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTGCCCCCTGCCCT 240

Db 162 CTTTGTGCACTCCACTGTGCAATGTGCGCCCTGCACGCTGGGGCTGTGCCCCCTGCCCT 103

Qy 241 TGGTCTGCCCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCTCAAT 300

Db 102 TGGTCTGCCCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCTCAAT 43

Qy 301 AAAGTGCACTGCTTGTGCAAAA 323

Db 42 AAAGTGCACTGCTTGTGCAAAA 20

RESULT 11

AD25685

ID AD25685 standard; cDNA; 1617 BP.

XX AC AD25685;

XX 29-JAN-2004 (first entry)

XX

DE Human cDNA differentially expressed in foam cells #89.
XX Human; ss; differential expression; foam cell; LPS; lipopolysaccharide;
KW cardiovascular disease; atherosclerosis.
XX
OS Homo sapiens.
XX US2003194721-A1.
XX 16-OCT-2003.
XX 18-SEP-2002; 2002US-00247671.
XX 19-SEP-2001; 2001US-0323784P.
XX (INCY-) INCYTE GENOMICS INC.
XX Mikita T, Shiffman D, Porter JG, Kaser MR;
XX WPI; 2003-875398/81.
XX
XX Combination containing several polynucleotide that are differentially
PT expressed in foam cells and complements of the polynucleotides, useful
PT for diagnosing cardiovascular disease or atherosclerosis.
XX
XX Claim 1; SEQ ID NO 89; 37pp; English.
XX
XX The invention relates to a combination comprising several polynucleotides
CC having any one of 127 sequences (S1) such as the sequence of human
CC calmodulin gene, human mRNA for KIAA0930 protein, leukotriene A4
CC hydrolase, human CGI-142 protein mRNA, human K+ channel beat 2 subunit
CC mRNA, etc., and their complements. The cDNAs are differentially expressed
CC in LPS (lipopolysaccharide)-treated foam cells. Also included are
CC obtaining an extended or full length gene from a library of nucleic acid
CC sequences, an expression vector containing the nucleic acids, a host cell
CC containing the vector, a purified polypeptide appearing as ADE25750 and
CC ADE25751, producing a protein by culturing the host cell, and a
CC composition comprising a purified antibody that specifically binds to the
CC proteins. The foam cell-expressed nucleic acids are useful for a high
CC throughput detection of differential expression of one or more
CC polynucleotides in a sample. The sample is from a subject with
CC atherosclerosis and comparison with a standard defines early, mid or late
CC stages of the disorder. The foam cell-expressed nucleic acids are useful
CC for high throughput screening of a library of molecules or compounds to
CC identify a ligand which binds a polynucleotide. The library is chosen
CC from DNA molecules, peptides, proteins and RNA molecules. The protein is
CC useful for a high throughput screening of library of molecules or
CC compounds to identify at least one ligand which specifically binds a
CC protein, for purifying a ligand from a sample for making an antibody. The
CC foam cell-expressed nucleic acids are useful for diagnosing
CC cardiovascular disorder. The foam cell-expressed nucleic acids are useful
CC as elements on a microarray which can be used for detecting related
CC polynucleotide in a sample, diagnosing cardiovascular disease,
CC atherosclerosis. The present sequence represents a cDNA whose expression
CC is upregulated in LPS treated foam cells.
XX
SQ Sequence 1617 BP; 348 A; 480 C; 445 G; 344 T; 0 U; 0 Other;

Query Match 98.3%; Score 317.4; DB 10; Length 1617;
Best Local Similarity 99.7%; Pred. No. 1.3e-89;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGGGCCAGGCTATGCTTCACCGCAAGTGCCTGAACATCTGGCCCTGATTCCTGGG 60
DB |||||||
773 CGGGGCCAGGCTATGCTTCACCGCAAGTGCCTGAACATCTGGCCCTGATTCCTGGG 832
QY 61 CATCCTCATGACCATTTGGATTACCCCTGTACTGGTATTGGCTCTGTGACAGTCTACCA 120
DB |||||||
833 CATCCTCATGACCATTTGGATTATCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 892
QY 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 180
DB |||||||
893 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 952

QY 181 CCTTTGACCTCCACTGTGCAATGCTGGCCCTGACGCTGGGGCTGTTGCCCTGCCCCCT 240
DB |||||||
953 CTTTGGACCTCCACTGTGCAATGCTGGCCCTGACGCTGGGGCTGTTGCCCTGCCCCCT 1012
QY 241 TGGTCTCTGCCCCCTAGATACAGCAGTATTATATACCCACACACCTGTCTACAGTGTCTTCAAT 300
DB |||||||
1013 TGGTCTCTGCCCCCTAGATACAGCAGTATTATATACCCACACACCTGTCTACAGTGTCTTCAAT 1072
QY 301 AAAGTGCACGTGCTTGTGA 319
DB |||||||
1073 AAAGTGCACGTGCTTGTGA 1091

RESULT 12
ABD32562
ID ABD32562 standard; DNA; 21478 BP.
XX
AC ABD32562;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human cancer-associated genomic DNA HD14-034.
XX
KW Human; ds; cancer-associated protein; gene; cytostatic; cancer;
KW leukaemia; lymphoma; CAP.
XX
OS Homo sapiens.
XX
PN WO2004074320-A2.
XX
PD 02-SEP-2004.
XX
PF 17-FEB-2004; 2004WO-US004730.
XX
PR 14-FEB-2003; 2003US-00367094.
PR 14-MAR-2003; 2003US-00388838.
PR 15-APR-2003; 2003US-00417375.
PR 13-JUN-2003; 2003US-00461862.
PR 15-SEP-2003; 2003US-00663431.
PR 15-DEC-2003; 2003US-00737318.
XX
XX (SAGR-) SAGRES DISCOVERY INC.
XX
XX Morris DW, Morris DW, Malandro MS;
PI WPI; 2004-652914/63.
XX
DR
XX
PT New isolated cancer-associated polynucleotides and polypeptides useful
PT for diagnosing, preventing or treating cancers, especially lymphoma and
PT leukemia, or in screening for agents that modulate cancer.
XX
XX claim 16; seqid 46; 310pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising at least 10
CC contiguous nucleotides of any of the 233 polynucleotide sequences given
CC in the specification, or its complement. The nucleic acids encode cancer-
CC associated proteins. Also included are an expression vector comprising
CC the isolated nucleic acid cited above, a host cell comprising the above
CC recombinant nucleic acid or expression vector, a microarray for detecting
CC a cancer-associated (CA) nucleic acid comprising at least one probe
CC comprising at least 10 contiguous nucleotides of any of the above-
CC mentioned nucleotide sequences, an isolated polypeptide (encoded within
CC an open reading frame of a CA sequence selected from any of the 95
CC polynucleotide sequences as mentioned in the specification, or its
CC complement), an isolated antibody, (or its antigen binding fragment) that
CC binds to the above polypeptide, a hybridoma that produces the above
CC monoclonal antibody, a pharmaceutical composition comprising the above
CC antibody and a pharmaceutical excipient, a kit for detecting cancer
CC cells (comprising the antibody cited above, methods for diagnosing cancer
CC or for detecting the presence or absence of cancer cells in an
CC individual, a method for inhibiting growth of cancer cells in an
CC individual, a method for delivering a therapeutic agent to cancer cells

CC in an individual, an electronic library comprising the above
CC polynucleotide or polypeptide (or their fragments), methods of screening
CC for anticancer activity or for a bioactive agent capable of modulating
CC the activity of a CA protein (CAP), methods for detecting cancer
CC associated with expression of a polypeptide in a test cell sample, a
CC method for treating cancers and a method for inhibiting the expression of
CC CA gene in a cell. The composition and methods are useful for detecting,
CC diagnosing, preventing and treating cancers, especially lymphoma and
CC leukaemia. These may also be used in screening for agents that modulate
CC cancer. The present sequence is a human CAP genomic sequence. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 21478 BP; 3963 A; 4821 C; 5071 G; 3826 T; 0 U; 3797 Other;

Query Match 98.3%; Score 317.4; DB 13; Length 21478;
Best Local Similarity 99.7%; Pred. No. 4.4e-89;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGGCCCCAGGCTTATGCTCCACCGCCAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
DB 17039 CGGGCCCCAGGCTTATGCTCCACCGCCAGTGCCTGAACATCTGGGCCCTGATTCGGG 17098
QY 61 CATCTCATGACCAATTGATTCACCCCTGTACTGTATTTCGGCTCTGTGACAGTCTACCA 120
DB 17099 CATCTCATGACCAATTGATTCATCTCTTACTGTATTTCGGCTCTGTGACAGTCTACCA 17158
QY 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180
DB 17159 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 17218
QY 181 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCT 240
DB 17219 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCT 17278
QY 241 TGGTCTCTGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCTTCAAT 300
DB 17279 TGGTCTCTGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCTTCAAT 17338
QY 301 AAAGTGCACTGCTTGTGA 319
DB 17339 AAAGTGCACTGCTTGTGA 17357

RESULT 13
ABQ60782
ID ABQ60782 standard; cDNA; 647 BP.
XX
AC ABQ60782;
XX
DT 02-AUG-2002 (first entry)
XX
DE Human 9-27 cDNA sequence SEQ ID NO:4484.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200229086-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US030732.
XX
PR 02-OCT-2000; 2000US-0237271P.
XX
PA (FARB) BAYER CORP.
XX
PI Burgess C, Astle JH, Carroll E, Catino TU, Dwivedi P, Molino GA;
PI Thiagalingam A, Lewis ME;

DR WPI; 2002-426115/45.
DR P-PSDB; ABB78999.
XX
PT New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell or
PT tissue type, and in antisense therapy.
XX
PS Claim 1; Fig 2; 796pp; English.
XX
CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists
XX
SQ Sequence 647 BP; 143 A; 211 C; 145 G; 148 T; 0 U; 0 Other;

Query Match 97.8%; Score 315.8; DB 6; Length 647;
Best Local Similarity 99.4%; Pred. No. 2.8e-89;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CGGGCCCCAGGCTTATGCTCCACCGCCAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
DB 329 CGGGCCCCAGGCTTATGCTCCACCGCCAGTGCCTGAACATCTGGGCCCTGATTCGGG 388
QY 61 CATCTCATGACCAATTGATTCACCCCTGTACTGTATTTCGGCTCTGTGACAGTCTACCA 120
DB 389 CATCTCATGACCAATTGATTCATCTCTTACTGTATTTCGGCTCTGTGACAGTCTACCA 448
QY 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180
DB 449 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 508
QY 181 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCT 240
DB 509 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCT 568
QY 241 TGGTCTCTGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCTTCAAT 300
DB 569 TGGTCTCTGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCTTCAAT 628
QY 301 AAAGTGCACTGCTTGTGA 319
DB 629 AAAGTGCACTGCTTGTGA 647

RESULT 14
ABX10351
ID ABX10351 standard; DNA; 647 BP.
XX
AC ABX10351;
XX
DT 28-JAN-2003 (first entry)
XX
DE DNA encoding protein differentially regulated in prostate cancer #20.
XX
KW Prostate cancer; gene expression; differential regulation;
KW molecular marker; drug target; cancer detection; cancer diagnosis;
KW cancer staging; cancer grading; cancer assessing; cancer monitoring;
XX
OS Homo sapiens.

```
XX WO200281638-A2.
XX PN 17-OCT-2002.
XX PD 08-APR-2002; 2002WO-US010824.
XX PF 06-APR-2001; 2001US-0281731P.
XX PG 06-APR-2001; 2001US-0281732P.
XX PH (ORIG-) ORIGENE TECHNOLOGIES INC.
XX PI Sun Z, Jay G;
XX P1 WPI; 2003-058520/05.
XX P2 P-PSDB; ABU07449.
XX DR
XX PT Novel genes which are differentially regulated in prostate cancer, useful
XX for diagnosing prostate cancer in prostate tissue sample and assessing
XX therapeutic or preventive intervention in prostate cancer patients.
XX PS Claim 1; Page 95; 416pp; English.
XX
XX The invention describes genes (I) which are differentially regulated in
XX prostate cancer. (I) is useful for diagnosing a prostate cancer in a
XX sample comprising prostate tissue, which involves determining the number
XX of target genes which are differentially-regulated in the sample, where
XX the number is indicative of the probability that the sample comprises
XX prostate cancer. (I) is useful for assessing a therapeutic or preventive
XX intervention in a subject having a prostate cancer, which involves
XX determining the expression levels in a sample comprising prostate tissue
XX of target genes which are differentially-regulated in prostate cancer.
XX Preferably, the expression levels of at least 10 genes are determined.
XX (I) is also useful for identifying agents that modulate a biological
XX activity of a polypeptide differentially-regulated in prostate cancer
XX cells, which involves contacting a polypeptide differentially-regulated
XX in prostate cancer cells with a test agent under conditions effective for
XX the test agent to modulate a biological activity of the polypeptide, and
XX determining whether the test agent modulates the biological activity. (I)
XX is useful as molecular markers, as drug targets, and for detecting,
XX diagnosing or treating, grading, assessing, monitoring, prognosticating,
XX conditions especially relating to prostate cancer. (I) and its expression
XX products are used in the diagnostic test to assay for presence of cancer
XX e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
XX blood etc. (I) is useful for assessing cancer e.g., to determine the type
XX of cancer, its stage of development, the nature of genetic defect, etc.
XX The polypeptide encoded by (I) can be used as target for therapy or drug
XX discovery. (I) can also be used for expressing the polypeptide and thus
XX for searching specific binding partners of the polypeptide. (I) is useful
XX in therapeutic applications to treat prostate cancer. The identification
XX of specific genes, and groups of genes, expressed in pathways
XX physiologically relevant to prostate cancer permits the definition of
XX functional and disease pathways and the delineation of targets in these
XX pathways which are useful in diagnostic, therapeutic, and clinical
XX applications. This sequence encodes a protein differentially regulated in
XX prostate cancer
XX SQ
XX
XX Query Match 97.8%; Score 315.8; DB 8; Length 647;
XX Best Local Similarity 99.4%; Pred. No. 2.8e-89;
XX Matches 31; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCCTGAAACATCTGGCCCTGATTCGGG 60
XX DB 329 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCCTGAAACATCTGGCCCTGATTCGGG 388
XX
XX QY 61 CATCCTCATGACCAATGGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120
XX DB 389 CATCCTCATGACCAATGGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 448
XX
XX QY 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 180
XX
XX Db 449 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 508
XX
XX QY 181 CTTTGCACCTCCACTGTGCAATGCTGGCCCTGACAGCTGGGGTGTGGCCCTGCCCCCT 240
XX
XX Db 509 CTTTGCACCTCCACTGTGCAATGCTGGCCCTGACAGCTGGGGTGTGGCCCTGCCCCCT 568
XX
XX QY 241 TGGTCTGCGCCCTAGATACAGCAGTATTATACCCACACACCTGCTACAGTGTCAATCAAT 300
XX
XX Db 569 TGGTCTGCGCCCTAGATACAGCAGTATTATACCCACACACCTGCTACAGTGTCAATCAAT 628
XX
XX QY 301 AAAGTGCACGTGCTTGTGA 319
XX
XX Db 629 AAAGTGCACGTGCTTGTGA 647
XX
XX RESULT 15
XX ADD18965
XX ID ADD18965 standard; DNA; 647 BP.
XX
XX AC ADD18965;
XX
XX DT 15-JAN-2004 (first entry)
XX
XX DE Human disease related protein DNA sequence SeqID454.
XX
XX KW human; disease state; cytostatic; antiinflammatory; ophthalmological;
XX antiarteriosclerotic; vulnary; gene therapy;
XX KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;
XX inflammation; erythropoiesis; glycolysis; gluconeogenesis;
XX glucose transportation; catecholamine synthesis; iron transport;
XX nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
XX retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
XX inflammatory condition; wound healing; gene; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO2003018621-A2.
XX
XX PD 06-MAR-2003.
XX
XX PF 23-AUG-2002; 2002WO-GB003892.
XX
XX PR 23-AUG-2001; 2001GB-00020558.
XX
XX PR 05-OCT-2001; 2001GB-00024037.
XX
XX PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX PI Kingman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
XX
XX DR WPI; 2003-230046/28.
XX
XX DR P-PSDB; ADD18964.
XX
XX PT New substantially purified polypeptide, useful for diagnosing or treating
XX a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
XX injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
XX wound healing.
XX
XX PS Claim 27; SEQ ID NO 454; 424pp; English.
XX
XX This invention relates to novel human genes and gene product which are
XX implicated in certain disease states. Compounds which modulate the
XX proteins of the invention may have cytostatic, antiinflammatory, the
XX ophthalmological, antiarteriosclerotic or vulnary activities. The
XX sequences of the invention may be useful for gene therapy. The invention
XX may be useful for diagnosing or treating a hypoxia-regulated condition,
XX such as tumorigenesis, angiogenesis, apoptosis, inflammation,
XX erythropoiesis, or the biological response to hypoxia conditions
XX including processes such as glycolysis, gluconeogenesis, glucose
XX transportation, catecholamine synthesis, iron transport or nitric oxide
XX synthesis. The disease includes cancer, ischaemic conditions, reperfusion
XX injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
XX inflammatory conditions or wound healing. The present sequence is that of
```



```
CC a disease related protein encoding DNA sequence of the invention.
XX
SQ Sequence 647 BP; 143 A; 211 C; 145 G; 148 T; 0 U; 0 Other;

Query Match          97.8%; Score 315.8; DB 10; Length 647;
Best Local Similarity 99.4%; Pred. No. 2.8e-89;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGGGCCCCAGGCTATGCTCCACCGCCAAGTGCCTGAACATCTGGGCCCTGATTTCTGGG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 329 CGGGCCCCAGGCTATGCTCCACCGCCAAGTGCCTGAACATCTGGGCCCTGATTTCTGGG 388
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 61 CATCTCATGACCATTTGGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 389 CATCTCATGACCATTTGGATTACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 448
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 121 TATTATGTTACAGATAATACAGGAACACGGGGTTACTAGTAGCGCCCATAGCCTGCAA 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 449 TATTATGTTACAGATAATACAGGAACACGGGGTTACTAGTAGCGCCCATAGCCTGCAA 508
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 181 CCTTTGCACTCCACTGTGCAATGCTTGGCCCTGCACTGGGCTGTTGCCCTTGCCCCCT 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 509 CCTTTGCACTCCACTGTGCAATGCTTGGCCCTGCACTGGGCTGTTGCCCTTGCCCCCT 568
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 241 TGGTCCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATTCAAT 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 569 TGGTCCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATTCAAT 628
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 301 AAAGTGCACGTGCTTGTA 319
Db ||||||||||||||||
Qy 629 AAAGTGCACGTGCTTGTA 647
Db ||||||||||||||||
```

Search completed: December 7, 2005, 00:03:07
Job time : 510.782 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:31:40 ; Search time 804.279 Seconds
(without alignments)
3320.997 Million cell updates/sec

Title: US-09-980-046B-6

Perfect score: 323

Sequence: 1 cggggccaggcctatgcttgcct.....gtgcacgtgcttggaataa 323

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main:

1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	319.8	99.0	1421	3	US-09-814-353-19857
2	319.4	98.9	1095	6	US-10-240-965-263
3	318.2	98.5	593	3	US-09-969-034-3967
4	317.4	98.3	1617	6	US-10-247-671-89
5	317.4	98.3	21478	7	US-10-367-094-201
6	315.8	97.8	647	3	US-09-969-034-4484
7	315.8	97.8	647	6	US-10-301-822-80
8	315.8	97.8	647	6	US-10-133-937-25
9	315.8	97.8	647	6	US-10-172-118-934
10	315.8	97.8	647	6	US-10-159-563-25
11	315.8	97.8	647	7	US-10-342-887-934
12	315.8	97.8	647	7	US-10-734-564-35
13	315.8	97.8	647	7	US-10-755-889-229
14	315.8	97.8	647	9	US-10-631-467-76
15	315.8	97.8	842	5	US-09-920-300A-1743
16	315.8	97.8	842	5	US-10-033-528-1743
17	315.8	97.8	842	6	US-10-099-926-1743
18	315.8	97.8	842	9	US-10-961-527-1743
19	315.8	97.8	853	3	US-09-880-107-2108
20	315.8	97.8	853	3	US-09-960-706-626
21	315.8	97.8	853	3	US-09-873-319-392
22	315.8	97.8	853	5	US-10-171-581-195
23	315.8	97.8	853	6	US-10-257-021-57

24	315.8	97.8	853	8	US-10-802-432-10	Sequence 10, Appl
25	315.8	97.8	853	9	US-10-505-680-52	Sequence 52, Appl
c 26	315.2	97.6	527	3	US-09-969-034-4116	Sequence 4116, Ap
c 27	311.8	96.5	764	3	US-09-810-936-283	Sequence 283, App
c 28	311.8	96.5	764	3	US-09-429-755-283	Sequence 283, App
c 29	311.8	96.5	764	3	US-09-924-400-283	Sequence 283, App
c 30	311.8	96.5	764	6	US-10-212-679-283	Sequence 283, App
c 31	311.8	96.5	764	7	US-10-079-137B-283	Sequence 283, App
c 32	311.6	96.5	519	3	US-09-969-034-4078	Sequence 4078, Ap
c 33	306.4	94.9	388	9	US-10-450-763-11579	Sequence 11579, A
34	306.2	94.8	1109	5	US-10-093-766-3	Sequence 3, Appli
35	306.2	94.8	1109	5	US-10-084-817-37	Sequence 37, Appl
36	306.2	94.8	1109	6	US-10-051-835-5	Sequence 5, Appli
c 37	305.4	94.6	852	5	US-09-814-353-14867	Sequence 27, Appl
c 38	302.2	93.6	427	3	US-10-176-847-27	Sequence 14867, A
c 39	293.2	90.8	430	9	US-10-505-680-53	Sequence 53, Appl
40	244.2	75.6	674	3	US-09-969-034-4224	Sequence 4224, Ap
41	219.8	68.0	651	8	US-10-128-558-116	Sequence 116, App
42	217.2	67.2	643	3	US-09-969-034-4364	Sequence 4364, Ap
c 43	210.8	65.3	345	3	US-09-814-353-14920	Sequence 14920, A
44	209.6	64.9	434	3	US-09-918-995-35636	Sequence 35636, A
45	208.6	64.6	579	3	US-09-969-034-4345	Sequence 4345, Ap

ALIGNMENTS

RESULT 1

US-09-814-353-19857
; Sequence 19857, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19857
; LENGTH: 1421
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-19857

Query Match 99.0%; Score 319.8; DB 3; Length 1421;
Best Local Similarity 99.4%; Pred. No. 1.7e-98;
Matches 321; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	CGGGGCCAGGCCCTATGCTCCACCGCAAGTGGCTGAACATCTGGGCCCTGATTCGGG	60
DB	420	CGGGGCCAGGCCCTATGCTCCACCGCAAGTGGCTGAACATCTGGGCCCTGATTCGGG	479
QY	61	CATCTCTATGACCAATGGATTCACCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA	120
DB	480	CATCTCTATGACCAATGGATTCATCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA	539

Qy 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180
Db 540 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 599
Qy 181 CTTTGGCACTCCACTGTGCAATGCTGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 240
Db 600 CTTTGGCACTCCACTGTGCAATGCTGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 659
Qy 241 TGGTCCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAAT 300
Db 660 TGGTCCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAAT 719
Qy 301 AAAGTGACAGTGTGTGAAAAA 323
Db 720 AAAGTGACAGTGTGTGAAAAA 742

RESULT 2
US-10-240-965-263
; Sequence 263, Application US/10240965
; Publication No. US20030165924A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: SHIFFMAN, Dov
; APPLICANT: SOMOGYI, Roland
; APPLICANT: LAWN, Richard M.
; APPLICANT: SEILHAMER, Jeffrey J.
; APPLICANT: PORTER, Gordon J.
; APPLICANT: MIKITA, Thomas
; APPLICANT: TAI, Julie
; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
; FILE REFERENCE: PA-0025 PCT
; CURRENT APPLICATION NUMBER: US/10/240,965
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 2002-10-04
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 276
; SOFTWARE: PERL Program
; SEQ ID NO 263
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030165924A1 988231.7
US-10-240-965-263

Query Match 98.9%; Score 319.4; DB 6; Length 1095;
Best Local Similarity 99.7%; Pred. No. 2.le-98;
Matches 320; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCCCTATGCTCCACCGCAAGTCTGAAACATCTGGGCCCTGATTTGGG 60
Db 775 CGGGGCCAGGCCCTATGCTCCACCGCAAGTCTGAAACATCTGGGCCCTGATTTGGG 834
Qy 61 CATCCTCATGACCATTTGATTCACCCCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 120
Db 835 CATCCTCATGACCATTTGATTCACCCCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 894
Qy 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180
Db 895 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 954
Qy 181 CTTTGGCACTCCACTGTGCAATGCTGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 240
Db 955 CTTTGGCACTCCACTGTGCAATGCTGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 1014
Qy 241 TGGTCCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAAT 300
Db 1015 TGGTCCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAAT 1074
Qy 301 AAAGTGACAGTGTGTGAAAA 321
Db 1015 TGGTCCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAAT 1074

Db 1075 AAAGTGACAGTGTGTGAAAA 1095

RESULT 3
US-09-969-034-3967/c
; Sequence 3967, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Aetle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poornima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 3967
; LENGTH: 593
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 380, 387, 407, 415, 424, 433, 436, 442, 467, 472, 493, 496,
; LOCATION: 520, 524, 533, 545, 548, 556, 559, 563
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-3967

Query Match 98.5%; Score 318.2; DB 3; Length 593;
Best Local Similarity 99.1%; Pred. No. 4.le-98;
Matches 320; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCCCTATGCTCCACCGCAAGTCTGAAACATCTGGGCCCTGATTTGGG 60
Db 342 CGGGGCCAGGCCCTATGCTCCACCGCAAGTCTGAAACATCTGGGCCCTGATTTGGG 283
Qy 61 CATCCTCATGACCATTTGATTCACCCCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 120
Db 282 CATCCTCATGACCATTTGATTCACCCCTGTTACTGGTATTGGCTCTGTGACAGTCTACCA 223
Qy 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180
Db 222 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 163
Qy 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 240
Db 162 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 103
Qy 241 TGGTCCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAAT 300
Db 102 TGGTCCCTGCCCTAGATACAGAGTTTATATACCCACACACCTGTCTACAGTGTCAAT 43
Qy 301 AAAGTGACAGTGTGTGAAAAA 323
Db 42 AAAGTGACAGTGTGTGAAAAA 20

RESULT 4
US-10-247-671-89
; Sequence 89, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.

APPLICANT: Kaser, Matthew R.
TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
FILE REFERENCE: PA-0050 US
CURRENT APPLICATION NUMBER: US/10/247,671
CURRENT FILING DATE: 2002-09-18
PRIOR APPLICATION NUMBER: 60/323,784
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 186
SOFTWARE: PERL Program
SEQ ID NO 89
LENGTH: 1617
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030194721A1 1328372.14
US-10-247-671-89

Query Match 98.3%; Score 317.4; DB 6; Length 1617;
Best Local Similarity 99.7%; Pred. No. 1.2e-97;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 60
DB 773 CGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 832
QY 61 CATCTCATGACCATTTGGATTCACCTCTGTACTGTATTTCGGCTCTGTGACAGTCTACCA 120
DB 833 CATCTCATGACCATTTGGATTCATCTGTACTGTATTTCGGCTCTGTGACAGTCTACCA 892
QY 121 TATTATGTTACAGATAAATACAGGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCCTGCAA 180
DB 893 TATTATGTTACAGATAAATACAGGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCCTGCAA 952
QY 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGACAGCTGGGGCTGTGGCCCTGCCCCCT 240
DB 953 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGACAGCTGGGGCTGTGGCCCTGCCCCCT 1012
QY 241 TGGTCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
DB 1013 TGGTCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 1072
QY 301 AAAGTGCACGTCTGTGTA 319
DB 1073 AAAGTGCACGTCTGTGTA 1091

RESULT 5
US-10-367-094-201
Sequence 201, Application US/10367094
Publication No. US20040170982A1
GENERAL INFORMATION:
APPLICANT: David W. Morris
APPLICANT: Marc Malandro
TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
FILE REFERENCE: 529452001500
CURRENT APPLICATION NUMBER: US/10/367,094
CURRENT FILING DATE: 2003-02-14
NUMBER OF SEQ ID NOS: 203
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 201
LENGTH: 21478
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(21478)
OTHER INFORMATION: n = A,T,C or G
US-10-367-094-201

Query Match 98.3%; Score 317.4; DB 7; Length 21478;
Best Local Similarity 99.7%; Pred. No. 3.9e-97;
Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 60
DB 17039 CGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 17098
QY 61 CATCTCATGACCATTTGGATTCACCTCTGTACTGTATTTCGGCTCTGTGACAGTCTACCA 120
DB 17099 CATCTCATGACCATTTGGATTCATCTCTGTACTGTATTTCGGCTCTGTGACAGTCTACCA 17158
QY 121 TATTATGTTACAGATAAATACAGGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCCTGCAA 180
DB 17159 TATTATGTTACAGATAAATACAGGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCCTGCAA 17218
QY 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGACAGCTGGGGCTGTGGCCCTGCCCCCT 240
DB 17219 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGACAGCTGGGGCTGTGGCCCTGCCCCCT 17278
QY 241 TGGTCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
DB 17279 TGGTCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 17338
QY 301 AAAGTGCACGTCTGTGTA 319
DB 17339 AAAGTGCACGTCTGTGTA 17357

RESULT 6
US-09-969-034-4484
Sequence 4484, Application US/09969034
Publication No. US20040110668A1
GENERAL INFORMATION:
APPLICANT: Burgess, Christopher C.
APPLICANT: Astle, Jon H.
APPLICANT: Carroll, Eddie III
APPLICANT: Catino, Theodore J.
APPLICANT: Dwivedi, Poorima
APPLICANT: Molino, Gary A.
APPLICANT: Thiagalingam, Arunthathi
APPLICANT: Lewis, Marcia E.
TITLE OF INVENTION: Nucleic Acid Sequences Differentially
Expressed in Cancer Tissue
FILE REFERENCE: 1657/1032
CURRENT APPLICATION NUMBER: US/09/969,034
CURRENT FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: 60/237,271
PRIOR FILING DATE: 2000-02-10
NUMBER OF SEQ ID NOS: 4494
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4484
LENGTH: 647
TYPE: DNA
ORGANISM: Homo sapiens
US-09-969-034-4484

Query Match 97.8%; Score 315.8; DB 3; Length 647;
Best Local Similarity 99.4%; Pred. No. 2.8e-97;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 60
DB 329 CGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCCTGGG 388
QY 61 CATCTCATGACCATTTGGATTCACCTCTGTACTGTATTTCGGCTCTGTGACAGTCTACCA 120
DB 389 CATCTCATGACCATTTGGATTCATCTCTGTACTGTATTTCGGCTCTGTGACAGTCTACCA 448
QY 121 TATTATGTTACAGATAAATACAGGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCCTGCAA 180
DB 449 TATTATGTTACAGATAAATACAGGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCCTGCAA 508
QY 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGACAGCTGGGGCTGTGGCCCTGCCCCCT 240
DB 509 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGACAGCTGGGGCTGTGGCCCTGCCCCCT 568

Qy 241 TGGTCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db |||||||
Qy 569 TGGTCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628
Db |||||||
Qy 301 AAAGTGCACGTGCTTGGA 319
Db |||||||
Qy 629 AAAGTGCACGTGCTTGGA 647

RESULT 7

US-10-301-822-80
; Sequence 80, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MP01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 647
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (111)...(488)
US-10-301-822-80

Query Match 97.8%; Score 315.8; DB 6; Length 647;
Best Local Similarity 99.4%; Pred. No. 2.8e-97;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGGGCCAGGCTATCGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATCTGGG 60
Db 329 CGGGGCCAGGCTATCGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATCTGGG 388
Qy 61 CATCCTCATGACCAATTGGATTACCCCTGTTACTGGTCTGTGACAGTCTACCA 120
Db 389 CATCCTCATGACCAATTGGATTACCCCTGTTACTGGTCTGTGACAGTCTACCA 448
Qy 121 TATTATGTTACAGATAACAGGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180
Db 449 TATTATGTTACAGATAACAGGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 508
Qy 181 CCTTTGCACCTGTCATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 509 CCTTTGCACCTGTCATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 568
Qy 241 TGGTCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 569 TGGTCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628
Qy 301 AAAGTGCACGTGCTTGGA 319
Db 629 AAAGTGCACGTGCTTGGA 647

RESULT 8

US-10-133-937-25
; Sequence 25, Application US/10133937
; Publication No. US20030207278A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING,
; TITLE OF INVENTION: DIAGNOSING, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND
; TITLE OF INVENTION: OTHER BIOLOGICAL STATES
; FILE REFERENCE: 11613.56US01
; CURRENT APPLICATION NUMBER: US/10/133,937
; CURRENT FILING DATE: 2002-11-04
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 647
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-133-937-25

Query Match 97.8%; Score 315.8; DB 6; Length 647;
Best Local Similarity 99.4%; Pred. No. 2.8e-97;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGGGCCAGGCTATCGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATCTGGG 60
Db 329 CGGGGCCAGGCTATCGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATCTGGG 388
Qy 61 CATCCTCATGACCAATTGGATTACCCCTGTTACTGGTCTGTGACAGTCTACCA 120
Db 389 CATCCTCATGACCAATTGGATTACCCCTGTTACTGGTCTGTGACAGTCTACCA 448
Qy 121 TATTATGTTACAGATAACAGGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180
Db 449 TATTATGTTACAGATAACAGGAAAAACGGGGTTACTAGTAGCCGCCCATAGCCTGCAA 508
Qy 181 CCTTTGCACCTGTCATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 509 CCTTTGCACCTGTCATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 568
Qy 241 TGGTCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 569 TGGTCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628
Qy 301 AAAGTGCACGTGCTTGGA 319
Db 629 AAAGTGCACGTGCTTGGA 647

RESULT 9

US-10-172-118-934
; Sequence 934, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veijer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14

```
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 934
; LENGTH: 647
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_003641
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-934

Query Match          97.8%; Score 315.8; DB 6; Length 647;
Best Local Similarity 99.4%; Pred. No. 2.8e-97;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAACTCTGGGCCCTGATTTCTGGG 60
DB 329 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAACTCTGGGCCCTGATTTCTGGG 388

QY 61 CATCTCATGACCATTTGCAATCACCTCTGTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
DB 389 CATCTCATGACCATTTGCAATCACCTCTGTACTGGTATTTCGGCTCTGTGAAAGTCTACCA 448

QY 121 TATTATGTTACAGATAATACAGGAAACCGGGGTTACTAGTAGCGGCCCATAGCTGCAA 180
DB 389 CATCTCATGACCATTTGCAATCACCTCTGTACTGGTATTTCGGCTCTGTGAAAGTCTACCA 448

QY 121 TATTATGTTACAGATAATACAGGAAACCGGGGTTACTAGTAGCGGCCCATAGCTGCAA 180
DB 449 TATTATGTTACAGATAATACAGGAAACCGGGGTTACTAGTAGCGGCCCATAGCTGCAA 508

QY 181 CTTTGGCACTCCACTGTCGAATGCTGCGCCTTGACCGCTGGGGCTGTTGGCCCTGCCCCCT 240
DB 509 CTTTGGCACTCCACTGTCGAATGCTGCGCCTTGACCGCTGGGGCTGTTGGCCCTGCCCCCT 568

QY 181 CTTTGGCACTCCACTGTCGAATGCTGCGCCTTGACCGCTGGGGCTGTTGGCCCTGCCCCCT 240
DB 509 CTTTGGCACTCCACTGTCGAATGCTGCGCCTTGACCGCTGGGGCTGTTGGCCCTGCCCCCT 568

QY 241 TGGTCTGCTGCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
DB 569 TGGTCTGCTGCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628

QY 301 AAAGTGCACGCTGCTTGTA 319
DB 629 AAAGTGCACGCTGCTTGTA 647

RESULT 10
US-10-159-563-25
; Sequence 25, Application US/10159563
; Publication No. US20040009154A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
; FILE REFERENCE: 11613.56US11
; CURRENT APPLICATION NUMBER: US/10/159,563
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 10/133,937
; PRIOR FILING DATE: 2002-04-25
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 647
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-159-563-25

Query Match          97.8%; Score 315.8; DB 6; Length 647;
Best Local Similarity 99.4%; Pred. No. 2.8e-97;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAACTCTGGGCCCTGATTTCTGGG 60
DB 329 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAACTCTGGGCCCTGATTTCTGGG 388

QY 61 CATCTCATGACCATTTGCAATCACCTCTGTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
DB 389 CATCTCATGACCATTTGCAATCACCTCTGTACTGGTATTTCGGCTCTGTGAAAGTCTACCA 448

QY 121 TATTATGTTACAGATAATACAGGAAACCGGGGTTACTAGTAGCGGCCCATAGCTGCAA 180
DB 449 TATTATGTTACAGATAATACAGGAAACCGGGGTTACTAGTAGCGGCCCATAGCTGCAA 508

QY 181 CTTTGGCACTCCACTGTCGAATGCTGCGCCTTGACCGCTGGGGCTGTTGGCCCTGCCCCCT 240
DB 509 CTTTGGCACTCCACTGTCGAATGCTGCGCCTTGACCGCTGGGGCTGTTGGCCCTGCCCCCT 568

QY 241 TGGTCTGCTGCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
DB 569 TGGTCTGCTGCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628

QY 301 AAAGTGCACGCTGCTTGTA 319
DB 629 AAAGTGCACGCTGCTTGTA 647

Query Match          97.8%; Score 315.8; DB 7; Length 647;
Best Local Similarity 99.4%; Pred. No. 2.8e-97;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAACTCTGGGCCCTGATTTCTGGG 60
DB 329 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAACTCTGGGCCCTGATTTCTGGG 388

QY 61 CATCTCATGACCATTTGCAATCACCTCTGTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
DB 389 CATCTCATGACCATTTGCAATCACCTCTGTACTGGTATTTCGGCTCTGTGAAAGTCTACCA 448

QY 121 TATTATGTTACAGATAATACAGGAAACCGGGGTTACTAGTAGCGGCCCATAGCTGCAA 180
DB 449 TATTATGTTACAGATAATACAGGAAACCGGGGTTACTAGTAGCGGCCCATAGCTGCAA 508

QY 181 CTTTGGCACTCCACTGTCGAATGCTGCGCCTTGACCGCTGGGGCTGTTGGCCCTGCCCCCT 240
DB 509 CTTTGGCACTCCACTGTCGAATGCTGCGCCTTGACCGCTGGGGCTGTTGGCCCTGCCCCCT 568

QY 241 TGGTCTGCTGCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
DB 569 TGGTCTGCTGCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628
```

Qy 301 AAATGACAGTGTGTGA 319
Db 629 AAATGACAGTGTGTGA 647

RESULT 12

US-10-734-564-35
; Sequence 35, Application US/10734564
; Publication No. US20040157278A1
; GENERAL INFORMATION:
; APPLICANT: Christopher C Burgess et al
; TITLE OF INVENTION: Detection Methods Using TIMP1
; FILE REFERENCE: 1657/2012
; CURRENT APPLICATION NUMBER: US/10/734,564
; CURRENT FILING DATE: 2003-12-12
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 647
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-734-564-35

Query Match 97.8%; Score 315.8; DB 7; Length 647;
Best Local Similarity 99.4%; Pred. No. 2.8e-97;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAAATCTGGGCCCTGATTTCTGGG 60
Db 329 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAAATCTGGGCCCTGATTTCTGGG 388
Qy 61 CATCTCATGACCAATGGATTACCCCTGATTCAGTCTGGTCTGTGACAGTCTACCA 120
Db 389 CATCTCATGACCAATGGATTACCCCTGATTCAGTCTGGTCTGTGACAGTCTACCA 448
Qy 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTACGCCCCATAGCCTGCAA 180
Db 449 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTACGCCCCATAGCCTGCAA 508
Qy 181 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240
Db 509 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 568
Qy 241 TGGTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 569 TGGTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628
Qy 301 AAATGACAGTGTGTGA 319
Db 629 AAATGACAGTGTGTGA 647

RESULT 13

US-10-755-889-229
; Sequence 229, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 229
; LENGTH: 647
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-755-889-229
Query Match 97.8%; Score 315.8; DB 7; Length 647;
Best Local Similarity 99.4%; Pred. No. 2.8e-97;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAAATCTGGGCCCTGATTTCTGGG 60
Db 329 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAAATCTGGGCCCTGATTTCTGGG 388
Qy 61 CATCTCATGACCAATGGATTACCCCTGATTCAGTCTGGTCTGTGACAGTCTACCA 120
Db 389 CATCTCATGACCAATGGATTACCCCTGATTCAGTCTGGTCTGTGACAGTCTACCA 448
Qy 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTACGCCCCATAGCCTGCAA 180
Db 449 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTACGCCCCATAGCCTGCAA 508
Qy 181 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240
Db 509 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 568
Qy 241 TGGTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 569 TGGTCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628
Qy 301 AAATGACAGTGTGTGA 319
Db 629 AAATGACAGTGTGTGA 647

RESULT 14

US-10-631-467-76
; Sequence 76, Application US/10631467
; Publication No. US20050208496A1
; GENERAL INFORMATION:
; APPLICANT: Genex Research Inc.
; TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive p
; FILE REFERENCE: 3462.1005-000
; CURRENT APPLICATION NUMBER: US/10/631,467
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: JP 2003-077212
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: JP 2002-229312
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2086
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76
; LENGTH: 647
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-631-467-76

Query Match 97.8%; Score 315.8; DB 9; Length 647;
Best Local Similarity 99.4%; Pred. No. 2.8e-97;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAAATCTGGGCCCTGATTTCTGGG 60
Db 329 CGGGGCCAGGCTATGCTCCACCGCAAGTGCTGAAATCTGGGCCCTGATTTCTGGG 388
Qy 61 CATCTCATGACCAATGGATTACCCCTGATTCAGTCTGGTCTGTGACAGTCTACCA 120
Db 389 CATCTCATGACCAATGGATTACCCCTGATTCAGTCTGGTCTGTGACAGTCTACCA 448
Qy 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTACGCCCCATAGCCTGCAA 180
Db 449 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTACGCCCCATAGCCTGCAA 508
Qy 181 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 240
Db 509 CCTTTGCACTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTTGCCCTGCCCCCT 568

Qy 241 TGGTCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 569 TGGTCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 628
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 301 AAAGTGCACGTGCTTGGA 319
Db ||||||||||||||||||||
Qy 629 AAAGTGCACGTGCTTGGA 647
Db ||||||||||||||||||||

RESULT 15
US-09-920-300A-1743
; Sequence 1743, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920.300A
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1743
; LENGTH: 842
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-920-300A-1743

Query Match 97.8%; Score 315.8; DB 3; Length 842;
Best Local Similarity 99.4%; Pred. No. 3.2e-97;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGGGCCAGGCCTATGCTCCACGGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 524 CGGGGCCAGGCCTATGCTCCACGGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 583
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 61 CATCCTCATGACCATTTGGATTACCCCTGTACTGGTATTCGGCTCTGTGACAGTCTACCA 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 584 CATCCTCATGACCATTTGGATTACCCCTGTACTGGTATTCGGCTCTGTGACAGTCTACCA 643
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 121 TATTATGTTACAGATAATACAGGAARACGGGTTACTAGTAGCGCCCATAGCCTGCAA 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 644 TATTATGTTACAGATAATACAGGAARACGGGTTACTAGTAGCGCCCATAGCCTGCAA 703
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 181 CCTTTGCACTCCACTGTGCAATGCTGGGCCCTGCAGCTGGGGCTGTTGCCCTGCCCCCT 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 704 CCTTTGCACTCCACTGTGCAATGCTGGGCCCTGCAGCTGGGGCTGTTGCCCTGCCCCCT 763
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 241 TGGTCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 764 TGGTCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 823
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 301 AAAGTGCACGTGCTTGGA 319
Db ||||||||||||||||||||
Qy 824 AAAGTGCACGTGCTTGGA 842
Db ||||||||||||||||||||

Search completed: December 6, 2005, 22:57:07
Job time : 806.279 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:34:11 ; Search time 220.487 Seconds
(without alignments)
455.919 Million cell updates/sec

Title: US-09-980-046B-6
Perfect score: 323
Sequence: 1 cggggccagcctatgcct.....gtgcagctgttgtaaaaa 323

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New:*
1: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB_seq.*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB_seq.*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB_seq.*
4: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB_seq.*
5: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB_seq.*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB_seq.*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB_seq.*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB_seq2.*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB_seq3.*
10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB_seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	116.2	36.0	182314	7	US-11-112-908-45
C 2	78.4	24.3	402	6	US-10-821-234-582
C 3	71	22.0	3667	6	US-10-750-185-25113
C 4	62.4	19.3	176	7	US-11-108-172-509
C 5	34	10.5	797	6	US-10-750-185-59126
C 6	32.6	10.1	3235	6	US-10-750-185-32680
C 7	31	9.6	3085	6	US-10-750-185-51883
C 8	30.4	9.4	1648	6	US-10-750-185-55588
C 9	30	9.3	1242	6	US-10-750-185-30919
C 10	29.8	9.2	1589	6	US-10-750-185-28916
C 11	29.8	9.2	1912	6	US-10-750-185-52845
C 12	29.4	9.1	744	6	US-10-750-185-39160
C 13	29.4	9.1	3648	6	US-10-750-185-35726
C 14	28.8	8.9	1053	6	US-10-750-185-47371
C 15	28.6	8.9	1250	6	US-10-750-185-19975
C 16	28.6	8.9	600	6	US-10-750-185-39178
C 17	28.6	8.9	1407	6	US-10-793-626-867
C 18	28.6	8.9	1407	6	US-10-793-626-1617
C 19	28.6	8.9	2563	6	US-10-750-185-35852
C 20	28.6	8.9	2580	6	US-10-793-626-4323
C 21	28.6	8.9	2856	6	US-10-793-626-3440
C 22	28.6	8.9	3227	6	US-10-793-626-3874
C 23	28.4	8.8	1605	6	US-10-750-185-53178

ALIGNMENTS

RESULT 1

US-11-112-908-45/c
; Sequence 45, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; PRIOR FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 45
; LENGTH: 182314
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-45

Query Match	36.0%;	Score 116.2;	DB 7;	Length 182314;
Best Local Similarity	66.0%;	Pred. No. 5.2e-28;		
Matches	208;	Conservative	0;	Mismatches 88;
Indels	19;	Gaps	2;	
Qy	13	CTATGCTCCACGCCCAAGTGCCTGAACATCTGGGCGCTGATCTGGGCATCTCTCATGAC	72	
Db	27380	CCATGCTTCCACTGCAAAAGTGCCTGAACATCTATGTCTCTGATCTTGGGCATCTCTGTAGC	27321	
Qy	73	CATTGGATTACCCCTGTACTGGTATTTCGGCTCTGTGACAGTCTACCATATATATTGTTACA	132	
Db	27320	CACGTGCATCATCATGATTTTACCACCATAGCTAGGGGATATCTCCCAAGCAATTTCTCA	27261	
Qy	133	GATAATACAGAAAAACGGGGTTACTAGTACCGCCCATAGCTGCAACCTTTTGCATCTCC	192	
Db	27260	GATGATTAAGGATCTCAAGGCCAGTAGTTG-----TGCCCATATATCTCC	27216	
Qy	193	ACTGTGCAATGCTGGCCCTGCAC-----GCTGGGCTGTGTGCCCCCTTGGTCTGTG	248	
Db	27215	ACCGTCCATTGCTGGCCTTGCACCTTGGCTGGGCTGTGTGGCCCTGCTGCTGCTGCTGCTG	27156	

QY	249	CCCTAGATACAGCAGTTTATATCCACACACCTGTCTCAGTGTGATCAATAAAGTCCA	308
Db	27155	CCCATTTATGCAGCAGTTTATCTACATACTTTCTCAAATGGTATTGATACAGGCA	27096
QY	309	CGTGCTTGCGAAAA	323
Db	27095	TGTGTTCTTGAAAAA	27081

```

RESULT 2
US-10-821-234-582
; Sequence 582, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt-seq_genes Version 1.0
; SEQ ID NO 582
; LENGTH: 402
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-582

```

	Query Match	24.3%	Score 78.4	DB 6	Length 402
	Best Local Similarity	84.6%	Pred. No. 1e-16		
	Matches 88	Conservative 0	Mismatches 16	Indels 0	Gaps 0
Qy	1	CGGGGCCAGGCTATGCTCTCACCGCGAAGTGCTGAA	CATCTGGGGCCTGATCTGGG	60	
Db	282	CGGGGCCAGGCTATGCTCTCACCGCGAAGTGCTGAA	CATCTGGGGCCTGATCTGGG	341	
Qy	61	CATCCTCATGACCAATTGGATCCACCTGTTACTGGTATTCGGCT		104	
Db	342	CATCCTCATGACCAATTGCTCATCGTCATCCCAAGTGTCTATCT		385	

```

RESULT 3
US-10-750-185-25113/c
; Sequence 25113, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25113
; LENGTH: 3667
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-25113

```

Query Match 22.0%; Score 71; DB 6; Length 3667;

	Best Local Similarity	65.4%;	Pred. No. 8.9e-14;					
	Matches	104;	Conservative	0;	Mismatches	55;	Indels	0; Gaps
Qy	2	GGGGCCAGCCCTT	GCCTCCACGGCCAGTGCTGAAACATCTGGGCCCTGATCTCTGGGC	61				
Db	1107	GGGGCCAGAGCT	TACGGCTCCACGGCAATGCGTGAACATCTGCTCCCTGGCTCTGGGC	1048				
Qy	62	ATCCTATGACCA	TGGATTACCCCTGTACTGGTATTCGGCTCTGTGACAGTCTACCAT	121				
Db	1047	ATCCTTCTGACT	CTGCTCTCATCGTCTGTCACCGGCTCCCTGATGATCGTTTCAG	988				
Qy	122	ATTATGTTACAT	ATAACAGGAAAAACGGGGTTACTAG	160				
Db	987	GCAGTCTCGGAG	CTCATGCAGAACTACGGAGGCCACTAG	949				

```

RESULT 4
US-11-108-172-509/c
; Sequence 509, Application US/11108172
; Publication No. US20050260177A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skelky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C15
; CURRENT APPLICATION NUMBER: US/11/108,172
; CURRENT FILING DATE: 2005-04-15
; PRIOR APPLICATION NUMBER: US 10/025,380
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US 09/922,217
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/833,263
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 09/649,811
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: US 09/609,448
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 09/575,251
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/519,444
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 09/504,629
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: US 09/480,321
; PRIOR FILING DATE: 2000-01-10
; PRIOR APPLICATION NUMBER: US 09/476,296
; PRIOR FILING DATE: 1999-12-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1130
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 509
; LENGTH: 176

```

Query Match 19.3%; Score 62.4; DB 7; Length 176;
Best Local Similarity 79.8%; Pred. No. 1.5e-11;

; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 55588
; LENGTH: 1648
; TYPE: DNA
; ORGANISM: Bovine 1986688068347
US-10-750-185-55588

Query Match 9.4%; Score 30.4; DB 6; Length 1648;
Best Local Similarity 59.1%; Pred. No. 2.1; Mismatches 0; Indels 0; Gaps 0;
Matches 52; Conservative 0

Qy 105 CTGTGACAGTCTACCATATATGTTACAGATAATACAGGAAACGGGGTTACTAGTAGC 164
Db 1406 CAGAGAGAGATGAATCTTCTATAAACTACAAATAAGGAAAGAGCTGTTAGAAAAAGC 1347

Qy 165 CGCCCATAGCTGCAACCTTTGCACTCC 192
Db 1346 AGTCTCTGCTGCAGCCTTTTCACCCC 1319

RESULT 9
US-10-750-185-30919
; Sequence 30919, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 30919
; LENGTH: 1242
; TYPE: DNA
; ORGANISM: Bovine 19866881328575
US-10-750-185-30919

Query Match 9.3%; Score 30; DB 6; Length 1242;
Best Local Similarity 48.8%; Pred. No. 2.5; Mismatches 0; Indels 0; Gaps 0;
Matches 81; Conservative 0

Qy 42 TCTGGCCCTGATCTCGGCATCTCTATGACCATTTGATTCACCTCTTACTGTATTCG 101
Db 1074 TCTGGAGCCTTGGCTGGGTTTCTACTGGGAAATTTGATGATGACTGGTCTACAAGATGG 1133

Qy 102 GCTCTGTGACAGTCTACCATATATGTTACAGATAATACAGGAAACGGGGTTACTAGT 161
Db 1134 GATCTGTATCTGTGAAGTGTGGCTCTTGACATACAGAGGATGTGGGATGCTCT 1193

Qy 162 AGCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGTCTGG 207
Db 1194 GGCTTCCCTCTTCTCTTGGCTCTTCTAGTCCCTCAGAACCTCTCG 1239

RESULT 10
US-10-750-185-28916/c
; Sequence 28916, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 28916
; LENGTH: 1589
; TYPE: DNA
; ORGANISM: Bovine 19866880854078
US-10-750-185-28916

Query Match 9.2%; Score 29.8; DB 6; Length 1589;
Best Local Similarity 55.2%; Pred. No. 3.3; Mismatches 0; Indels 0; Gaps 0;
Matches 58; Conservative 0

Qy 156 ACTAGTAGCGCCCATAGCTGCAACCTTTGCACTCCACTGTGCAATGCTGGCCCTGCAC 215
Db 1588 ACTTTTAGGAGCCTGAATCCAGGAGACAGAAAGCCCTCAGTGTGTCATGACCTGCC 1529

Qy 216 GCTGGGGCTGTTCGCCCTGCCCTTGGTCTCGCTCCCTAGATACA 260
Db 1528 CGTATGGCTGCTTCCTCCGAGACCTGGTTCGTCCTCACCCTTCA 1484

RESULT 11
US-10-750-185-52845
; Sequence 52845, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 52845
; LENGTH: 1912
; TYPE: DNA
; ORGANISM: Bovine 19866880622237
US-10-750-185-52845

Query Match 9.2%; Score 29.8; DB 6; Length 1912;
Best Local Similarity 52.9%; Pred. No. 3.6; Mismatches 0; Indels 0; Gaps 0;
Matches 64; Conservative 0

Qy 181 CCTTTGCACTCCCACTGTGCAATGCTGGCCCTTGACCGCTGGGGCTGTGGCCCTGCCCT 240
Db 1572 CCTTCTCTCTCTCTTGAGAATCTTGACCTTATCTCTCTTGAGAGCCCGACACCCCT 1631

Qy 241 TGGTCTGCCCCCTAGATACAGCAGTTTATACCACACACCTGTCTACAGTGTCTCAAT 300
Db 1632 TTCTCTCTCTCAGTGACCTGGACTTTTCTCAAGCTGCCTAGGAGTGAATCTCTCAAT 1691

Qy 301 A 301

```
Db      1692 A 1692

RESULT 12
US-10-750-185-39160
; Sequence 39160, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39160
; LENGTH: 744
; TYPE: DNA
; ORGANISM: Bovine 19866881456696
US-10-750-185-39160

Query Match      9.1%; Score 29.4; DB 6; Length 744;
Best Local Similarity 56.8%; Pred. No. 3;
Matches 54; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy      4 GGCCAGCCCTATGCTCCACGCCAAGTGCCTGAACATCTGGGCCCTGATCTGGGGCAT 63
Db      30 GTCTCGGACTGTGTCAGCCAAAGATATCGAATGGACTGCCATGCCCTCTCCAGGGCAT 89

Qy      64 CCTCATGACCATGGATTCAACCTGTTACTGGTAT 98
Db      90 CTTCCTGACCCAGAGATCAAAACCCGTCCTCTAT 124

RESULT 13
US-10-750-185-35726
; Sequence 35726, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35726
; LENGTH: 3648
; TYPE: DNA
; ORGANISM: Bovine 19866880517544
US-10-750-185-35726

Query Match      9.1%; Score 29.4; DB 6; Length 3648;
Best Local Similarity 52.9%; Pred. No. 6.7;
Matches 63; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Db      1692 A 1692

Qy      42 TCTGGGCCCTGATCTCTGGGCATCTCATGACCATTTGGATTCACCCCTGTTACTGTGATTCG 101
Db      2749 TCTATGCCCTCTCCAGGGGATCTTCTGACCCAGGGATTTGAACCCAGGTCTCCACATT 2808

Qy      102 GCTCTGTGACAGTCTACCATATTTATTGTACAGATATACAGGAAAAACGGGTTACTAG 160
Db      2809 GCAGGTGGATTCTTTTACCATCTGAGCCACCAGGGAAACCCAAAATATCTGGAGTGGGTAG 2867

RESULT 14
US-10-750-185-47371
; Sequence 47371, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47371
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Bovine 19866880909469
US-10-750-185-47371

Query Match      8.9%; Score 28.8; DB 6; Length 1053;
Best Local Similarity 58.0%; Pred. No. 5.7;
Matches 51; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy      10 GGCCATGCTCTCCACGCCAAGTGCCTGAACATCTGGGCCCTGATCTGGGCATCCCAT 69
Db      936 GGCTTCTCCAGGCAATAATATCTGGAGTGGATCACCAGGCCCTGCTTCAGGGGATCTTCCT 995

Qy      70 GACCATTTGGATTCAACCTGTTACTGGTA 97
Db      996 GACTCAGGGATTGAACCTGCAACTCCTA 1023

RESULT 15
US-10-750-185-19975
; Sequence 19975, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19975
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT04626
US-10-750-185-19975
```

Query Match 8.9%; Score 28.6; DB 6; Length 600;
Best Local Similarity 64.2%; Pred. No. 5.1; 24; Indels 0; Gaps 0;
Matches 43; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 103 CTCTGTGACAGTCTTACCATATTTATGTTACAGATAATACAGGAAAAACGGGTTACTAGTA 162
Db 150 CTGTGGGACAGTCTTCGCTTCCACGAAACAGAGAGAGGAGGAAAAATTCATTACTGGAT 209
Qy 163 GCCGCC 169
Db 210 GCTGCC 216

Search completed: December 6, 2005, 23:16:52
Job time : 221.487 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 3686.8 Seconds
(without alignments)
4099.008 Million cell updates/sec

Title: US-09-980-046B-6
Perfect score: 323
Sequence: 1 cggggccaggcctatgcct.....gtgcagtgcttgtaaaaa 323

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hcc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_ges1:*
10: gb_ges2:*
11: gb_ges3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	321.4	99.5	365	2	BG194579 RST13741
2	321.4	99.5	365	2	BG218402 RST18270
3	321.4	99.5	366	2	BG218401 RST18269
4	321.4	99.5	367	2	BG186275 RST5236 A
5	321.4	99.5	367	2	BG188986 RST7920 A
6	321.4	99.5	367	2	BG203761 RST23136
7	321.4	99.5	370	2	BG192020 RST11127
8	321.4	99.5	414	3	BM662200 UI-E-C10-
9	321.4	99.5	469	7	CN264982 170005999
10	321.4	99.5	498	1	AW574932 UI-HF-BL0
11	321.4	99.5	498	6	C3420737 UI-H-FG0-
12	321.4	99.5	504	5	BQ447972 UI-H-EU1-
13	321.4	99.5	505	3	BM672499 UI-E-CQ1-
14	321.4	99.5	510	3	BM767351 K-EST0049
15	321.4	99.5	522	5	BU783190 in04d02.x
16	321.4	99.5	552	6	CA414261 UI-H-EZ0-
17	321.4	99.5	563	7	CT001900 CT001900
18	321.4	99.5	567	5	BQ776064 UI-H-FR0-
19	321.4	99.5	568	5	BQ706056 AGENCOURT
20	321.4	99.5	574	5	BU731878 UI-E-C11-
21	321.4	99.5	587	5	BU680316 UI-CF-DU1
22	321.4	99.5	588	5	BU731736 UI-E-C11-

C 23	321.4	99.5	589	6	CA413789
C 24	321.4	99.5	611	6	CA428144
C 25	321.4	99.5	614	5	BQ440412
26	321.4	99.5	629	7	CR983622
27	321.4	99.5	635	7	CR985541
28	321.4	99.5	641	7	CR982834
C 29	321.4	99.5	645	6	CA414705
C 30	321.4	99.5	649	3	BQ181165
31	321.4	99.5	651	8	DR155863
32	321.4	99.5	652	8	DR158870
C 33	321.4	99.5	654	3	BQ181609
C 34	321.4	99.5	655	8	DR159276
C 35	321.4	99.5	656	5	BU686429
C 36	321.4	99.5	670	3	BM995109
37	321.4	99.5	674	7	CN264991
38	321.4	99.5	675	8	DR158734
C 39	321.4	99.5	677	6	CA421204
C 40	321.4	99.5	683	6	CA419982
41	321.4	99.5	683	8	CX165582
42	321.4	99.5	691	3	BQ219055
43	321.4	99.5	704	6	CD173064
C 44	321.4	99.5	706	6	CA429526
45	321.4	99.5	710	6	CD172976

ALIGNMENTS

RESULT 1
BG194579
LOCUS
DEFINITION RST13741 Athersys RAGE Library Homo sapiens cdna, mRNA linear EST 21-APR-2001
ACCESSION BG194579
VERSION BG194579.1 GI:13716266
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 365)
AUTHORS Harrington,J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
Mays,R., Smith,R., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
Offenbacher,J., Danzig,J. and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random
activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
PUBMED 11329013
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 365.
Location/Qualifiers
1. .365
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 99.5%; Score 321.4; DB 2; Length 365;

Best Local Similarity 99.7%; Pred. No. 1.3e-81;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCCTGGG 60
Db 35 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCCTGGG 94

Qy 61 CATCCTCATGACCAATTCGATTCACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120
Db 95 CATCCTCATGACCAATTCGATTCACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 154

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180
Db 155 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 214

Qy 181 CTTTGGCACTCACTGTGCAATGCTGGCCCTGCAAGTGGGGCTGTGGCCCTGCCCCCT 240
Db 215 CTTTGGCACTCACTGTGCAATGCTGGCCCTGCAAGTGGGGCTGTGGCCCTGCCCCCT 274

Qy 241 TGGTCCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 275 TGGTCCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 334

Qy 301 AAAGTGCACGTGCTTGTAAGAAA 323
Db 335 AAAGTGCACGTGCTTGTAAGAAA 357

RESULT 2
BG218402
LOCUS BG218402 365 bp mRNA linear EST 21-APR-2001
DEFINITION RST38270 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG218402
VERSION BG218402.1 GI:13744551
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 365)
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
Mays,R., Smith,B., Veloso,N., Klika,A., Hess,J., Cochren,K., Lo,K.,
Offenbacher,J., Danzig,J. and Ducar,M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
11329013
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athersys.com
High quality sequence stop: 365.
Location/Qualifiers
FEATURES
source
1. .365
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

Query Match 99.5%; Score 321.4; DB 2; Length 365;

ORIGIN

Best Local Similarity 99.7%; Pred. No. 1.3e-81;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCCTGGG 60
Db 35 CGGGGCCAGCCCTATGCTCCACCGCAAGTGCTGAAACATCTGGGCCCTGATTCCTGGG 94

Qy 61 CATCCTCATGACCAATTCGATTCACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 120
Db 95 CATCCTCATGACCAATTCGATTCACCCCTGTTACTGGTATTCGGCTCTGTGACAGTCTACCA 154

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 180
Db 155 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAA 214

Qy 181 CTTTGGCACTCACTGTGCAATGCTGGCCCTGCAAGTGGGGCTGTGGCCCTGCCCCCT 240
Db 215 CTTTGGCACTCACTGTGCAATGCTGGCCCTGCAAGTGGGGCTGTGGCCCTGCCCCCT 274

Qy 241 TGGTCCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 275 TGGTCCCTGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 334

Qy 301 AAAGTGCACGTGCTTGTAAGAAA 323
Db 335 AAAGTGCACGTGCTTGTAAGAAA 357

RESULT 3
BG218401
LOCUS BG218401 366 bp mRNA linear EST 21-APR-2001
DEFINITION RST38269 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG218401
VERSION BG218401.1 GI:13744550
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 366)
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
Mays,R., Smith,B., Veloso,N., Klika,A., Hess,J., Cochren,K., Lo,K.,
Offenbacher,J., Danzig,J. and Ducar,M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
11329013
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athersys.com
High quality sequence stop: 366.
Location/Qualifiers
FEATURES
source
1. .366
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="HT1080"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

Query Match 99.5%; Score 321.4; DB 2; Length 366;

ORIGIN

Best Local Similarity 99.7%; Pred. No. 1.3e-81;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGGGCCAGGCTATGCTCCACGGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db 36 CGGGGCCAGGCTATGCTCCACGGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 95

QY 61 CATCCTCATGACCAATGGATTCAACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
Db 96 CATCCTCATGACCAATGGATTCAACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 155

QY 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAG 180
Db 156 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAG 215

QY 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCGACGCTGGGGCTGTGGCCCTGCCCT 240
Db 216 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCGACGCTGGGGCTGTGGCCCTGCCCT 275

QY 241 TGGTCTGCCCTAGATACAGCAGTTTATATACCCACACACACCTGTCTACAGTGTCAATCAAT 300
Db 276 TGGTCTGCCCTAGATACAGCAGTTTATATACCCACACACACCTGTCTACAGTGTCAATCAAT 335

QY 301 AAAGTGCACGTGCTTGTAAGAAA 323
Db 336 AAAGTGCACGTGCTTGTAAGAAA 358

RESULT 4
BG186275
LOCUS
DEFINITION
RST5236 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION
BG186275
VERSION
BG186275.1 GI:13707962
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 367)
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
Offenbacher,J., Danzig,J. and Ducar,M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
11329013
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 367.
Location/Qualifiers
1..367
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 99.5%; Score 321.4; DB 2; Length 367;

Best Local Similarity 99.7%; Pred. No. 1.3e-81;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGGGCCAGGCTATGCTCCACGGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db 37 CGGGGCCAGGCTATGCTCCACGGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 96

QY 61 CATCCTCATGACCAATGGATTCAACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
Db 97 CATCCTCATGACCAATGGATTCAACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 156

QY 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAG 180
Db 157 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCCTGCAG 216

QY 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCGACGCTGGGGCTGTGGCCCTGCCCT 240
Db 217 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCGACGCTGGGGCTGTGGCCCTGCCCT 276

QY 241 TGGTCTGCCCTAGATACAGCAGTTTATATACCCACACACACCTGTCTACAGTGTCAATCAAT 300
Db 277 TGGTCTGCCCTAGATACAGCAGTTTATATACCCACACACACCTGTCTACAGTGTCAATCAAT 336

QY 301 AAAGTGCACGTGCTTGTAAGAAA 323
Db 337 AAAGTGCACGTGCTTGTAAGAAA 359

RESULT 5
BG188886
LOCUS
DEFINITION
RST7920 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION
BG188886
VERSION
BG188886.1 GI:13710573
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 367)
Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
Offenbacher,J., Danzig,J. and Ducar,M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
11329013
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 367.
Location/Qualifiers
1..367
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 99.5%; Score 321.4; DB 2; Length 367;

Best Local Similarity 99.7%; Pred. No. 1.3e-81;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGCCCTATGCTCCACCGCCAAAGTGCTGAAACATCTGGGCCCTGATTCTGGG 60
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
37 CGGGGCCAGCCCTATGCTCCACCGCCAAAGTGCTGAAACATCTGGGCCCTGATTCTGGG 96
Qy 61 CATCCTCATGACCAATTCGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
97 CATCCTCATGACCAATTCGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 156
Qy 121 TATTATGTTACAGATAATACAGGAAAAACCGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
157 TATTATGTTACAGATATACAGGAAAAACCGGGTTACTAGTAGCCGCCCATAGCCTGCAA 216
Qy 181 CTTTGTGACCTCCACTGTGCAATGTGCGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 240
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
217 CTTTGTGACCTCCACTGTGCAATGTGCGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 276
Qy 241 TGGTCTGCGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
277 TGGTCTGCGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 336
Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
337 AAAGTGCACGTGCTTGTGAAAAA 359

RESULT 6
BG192020
LOCUS BG192020 367 bp mRNA linear EST 21-APR-2001
DEFINITION RST23136 Atherys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG203761
VERSION BG203761.1 GI:13725432
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 367)
AUTHORS Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,
Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,
Mays, R., Smith, S., Veloso, N., Klika, A., Hess, J., Cochren, K., Lo, K.,
Offenbacher, J., Danzig, J. and Ducar, M.
TITLE Creation of genome-wide protein expression libraries using random
activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
PUBMED 11329013
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 367.
Location/Qualifiers
FEATURES
source
1. .367
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="HT1080"
/notes="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

Query Match 99.5%; Score 321.4; DB 2; Length 367;

ORIGIN

Best Local Similarity 99.7%; Pred. No. 1.3e-81;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGCCCTATGCTCCACCGCCAAAGTGCTGAAACATCTGGGCCCTGATTCTGGG 60
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
37 CGGGGCCAGCCCTATGCTCCACCGCCAAAGTGCTGAAACATCTGGGCCCTGATTCTGGG 96
Qy 61 CATCCTCATGACCAATTCGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
97 CATCCTCATGACCAATTCGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 156
Qy 121 TATTATGTTACAGATAATACAGGAAAAACCGGGTTACTAGTAGCCGCCCATAGCCTGCAA 180
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
157 TATTATGTTACAGATATACAGGAAAAACCGGGTTACTAGTAGCCGCCCATAGCCTGCAA 216
Qy 181 CTTTGTGACCTCCACTGTGCAATGTGCGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 240
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
217 CTTTGTGACCTCCACTGTGCAATGTGCGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 276
Qy 241 TGGTCTGCGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
277 TGGTCTGCGCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 336
Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
337 AAAGTGCACGTGCTTGTGAAAAA 359

RESULT 7
BG192020
LOCUS BG192020 370 bp mRNA linear EST 21-APR-2001
DEFINITION RST11127 Atherys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG192020
VERSION BG192020.1 GI:13713707
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 370)
AUTHORS Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,
Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,
Mays, R., Smith, S., Veloso, N., Klika, A., Hess, J., Cochren, K., Lo, K.,
Offenbacher, J., Danzig, J. and Ducar, M.
TITLE Creation of genome-wide protein expression libraries using random
activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
PUBMED 11329013
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 370.
Location/Qualifiers
FEATURES
source
1. .370
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="HT1080"
/notes="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

Query Match 99.5%; Score 321.4; DB 2; Length 370;

ORIGIN

Best Local Similarity 99.7%; Pred. No. 1.3e-81;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTTATGCTCCACGGCCAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db 40 CGGGGCCAGGCTTATGCTCCACGGCCAGTGCCTGAACATCTGGGCCCTGATTCGGG 99

Qy 61 CATCTCATGACCATTTGGATTCACTGTTACTGCTATTTCCGGCTCTGTGACAGTCTACCA 120
Db 100 CATCTCATGACCATTTGGATTCACTGTTACTGCTATTTCCGGCTCTGTGACAGTCTACCA 159

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 180
Db 160 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 219

Qy 181 CCTTTGCACTCCACTGTCATGCTGGCCCTGACGCTGGGGCTGTGGCCCTGCCCT 240
Db 220 CCTTTGCACTCCACTGTCATGCTGGCCCTGACGCTGGGGCTGTGGCCCTGCCCT 279

Qy 241 TGGTCTGCCCTTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 280 TGGTCTGCCCTTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 339

Qy 301 AAAGTGCACTGCTGTGTGAAAAA 323
Db 340 AAAGTGCACTGCTGTGTGAAAAA 362

RESULT 8
BM662200/c
LOCUS
DEFINITION
UI-E-CL0-acd-g-01-0-UI-s1 UI-E-CL0 Homo sapiens cDNA clone
UI-E-CL0-acd-g-01-0-UI 3', mRNA sequence.
ACCESSION
BM662200
VERSION
BM662200.1 GI:18966548
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 414)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
PUBMED
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRP, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Forward
POLYA=Yes.
Location/Qualifiers
1..414
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-CL0-acd-g-01-0-UI"
/tissue_type="human retina"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CL0"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-CL0 is a cDNA library containing the following
tissue(s): retina. The library was constructed according
to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT7T3-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CCGG. This library was
created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI).
TAG TISSUE=human retina
TAG_L18=UI-E-CL0
TAG_SEQ=CCGG

ORIGIN

Query Match 99.5%; Score 321.4; DB 3; Length 414;
Best Local Similarity 99.7%; Pred. No. 1.3e-81;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTTATGCTCCACGGCCAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db 336 CGGGGCCAGGCTTATGCTCCACGGCCAGTGCCTGAACATCTGGGCCCTGATTCGGG 277

Qy 61 CATCTCATGACCATTTGGATTCACTGTTACTGCTATTTCCGGCTCTGTGACAGTCTACCA 120
Db 276 CATCTCATGACCATTTGGATTCACTGTTACTGCTATTTCCGGCTCTGTGACAGTCTACCA 217

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 180
Db 216 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGGCCCATAGCTGCAA 157

Qy 181 CCTTTGCACTCCACTGTCATGCTGGCCCTGACGCTGGGGCTGTGGCCCTGCCCT 240
Db 156 CCTTTGCACTCCACTGTCATGCTGGCCCTGACGCTGGGGCTGTGGCCCTGCCCT 97

Qy 241 TGGTCTGCCCTTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300
Db 96 TGGTCTGCCCTTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 37

Qy 301 AAAGTGCACTGCTGTGTGAAAAA 323
Db 36 AAAGTGCACTGCTGTGTGAAAAA 14

RESULT 9
CN264982
LOCUS
DEFINITION
17000599934645 GRN_PRENEU Homo sapiens cDNA 5', mRNA sequence.
ACCESSION
CN264982
VERSION
CN264982.1 GI:47281396
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 469)
Bradenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
15146197
Contact: Bradenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658

Fax: 650 473 7760
 Email: rbrandenberger@eron.com
 Insert Length: 469 Std Error: 0.00.
 Location/Qualifiers

FEATURES

source

1. .469
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cell, retinoic acid and
 mitogen-treated hES cell line H7"
 /clone_lib="GRN PRENEU"
 /notes="oligo dt primed, full-length enriched cDNA library
 from hES cell line H7 (p29) maintained in feeder-free
 conditions. Embryoid bodies were generated in the presence
 of all-trans retinoic acid and mitogens."

ORIGIN

Query Match 99.5%; Score 321.4; DB 7; Length 469;
 Best Local Similarity 99.7%; Pred. No. 1.3e-81;
 Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGCCCTGATTCGGG 60
 DB 142 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGCCCTGATTCGGG 201
 QY 61 CATCCTCATGACCAATTCGATTCACCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
 DB 202 CATCCTCATGACCAATTCGATTCACCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 261
 QY 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTACCGCCCATAGCTGCA 180
 DB 262 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTACCGCCCATAGCTGCA 331
 QY 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 240
 DB 322 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 381
 QY 241 TGGTCTCGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
 DB 382 TGGTCTCGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 441
 QY 301 AAGTGCACGTGCTTGTGAAAAA 323
 DB 442 AAGTGCACGTGCTTGTGAAAAA 464

RESULT 10
 AW574932/c
 LOCUS
 DEFINITION UI-HF-BLO-abp-g-08-0-UI.s1 NIH_MGC_37 Homo sapiens cDNA clone
 IMAGE:3057590 3', mRNA sequence.

AW574932.1 GI:7246471

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov

The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: M.B. Soares Lab
 cDNA Library Arrayed by: M.B. Soares Lab
 DNA Sequencing by: M.B. Soares Lab
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: M13 Forward
 POLYA=Yes.

FEATURES

source

Location/Qualifiers
 1. .498
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:3057590"
 /tissue_type="lymph"
 /cell_type="germinal center B cells"
 /cell_line="MGC85"
 /lab_host="DH10B (LTI)"
 /clone_lib="NIH_MGC_37"
 /notes="Vector: pT7T3-Pac; Site 1: NotI; Site 2: Eco RI;
 Constructed from size fractionated cytoplasmic mRNA
 (1.5-2.5kb). Directionally cloned. Cells provided by Louis
 M. Staudt, Ph.D. Library preparation by Maria de Fatima
 Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

ORIGIN

Query Match 99.5%; Score 321.4; DB 1; Length 498;
 Best Local Similarity 99.7%; Pred. No. 1.4e-81;
 Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGCCCTGATTCGGG 60
 DB 335 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGCCCTGATTCGGG 276
 QY 61 CATCCTCATGACCAATTCGATTCACCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
 DB 275 CATCCTCATGACCAATTCGATTCACCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 216
 QY 121 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTACCGCCCATAGCTGCAA 180
 DB 215 TATTATGTTACAGATAATACAGAAAAACGGGGTTACTAGTACCGCCCATAGCTGCAA 156
 QY 181 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 240
 DB 155 CTTTGGCACTCCACTGTGCAATGCTGGCCCTGCAAGCTGGGGCTGTGGCCCTGCCCCCT 96
 QY 241 TGGTCTCGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 300
 DB 95 TGGTCTCGCCCTAGATACAGCAGTTTATACCCACACACCTGTCTACAGTGTCAATCAAT 36
 QY 301 AAGTGCACGTGCTTGTGAAAAA 323
 DB 35 AAGTGCACGTGCTTGTGAAAAA 13

RESULT 11
 CA420737/c

LOCUS
 DEFINITION UI-H-FG0-bc-q-b-23-0-UI.s1 NCI CGAP EN1 2 Homo sapiens cDNA clone
 UI-H-FG0-bc-q-b-23-0-UI 3', mRNA sequence.

CA420737.1 GI:24783392

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: James Martin
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source

Location/Qualifiers
1..498
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FG0-bcq-b-23-0-UI"
/tissue_type="Enchondroma cell line"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP EN1_2"
/note="Organ: Bone; Vector: p7773-Pac (Pharmacia) with a
modified polylinker; Site1: EcoR I; Site2: Not I;
NCI CGAP EN1_2 is a cDNA library containing the following
tissue(s): Enchondroma cell line (2 cell lines). The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into p7773-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is CCGGTCACTC. The cell lines was provided by Dr
James Martin from University of Iowa.
TAG_TISSUE=Enchondroma cell line (Mix of EN1 and EN2)
TAG_LIB=UI-H-FG0
TAG_SEQ=CGGTCACTC

ORIGIN

Query Match 99.5%; Score 321.4; DB 6; Length 498;
Best Local Similarity 99.7%; Pred. No. 1.4e-81;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db 336 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 277
Qy 61 CATCCTCATGACCATTTGATTCACCTCTTACTGTTACTGGTCTGTGACAGTCTACCA 120
Db 276 CATCCTCATGACCATTTGATTCACCTCTTACTGTTACTGGTCTGTGACAGTCTACCA 217
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCTGCAA 180
Db 216 TATTATGTTACAGATAATACAGGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCTGCAA 157
Qy 181 CTTTGGCACTCAGTGCATGTCGGCCCTGACGCTGGGGCTGTTGCCCTGCCCCCT 240
Db 156 CTTTGGCACTCAGTGCATGTCGGCCCTGACGCTGGGGCTGTTGCCCTGCCCCCT 97
Qy 241 TGGTCTGCCCTAGATACAGGATTTATACCCACACCTGTCTACAGTGTCAATCAAT 300
Db 96 TGGTCTGCCCTAGATACAGGATTTATACCCACACCTGTCTACAGTGTCAATCAAT 37
Qy 301 AAGTGCACGTGCTTGTGAAAA 323
Db 36 AAGTGCACGTGCTTGTGAAAA 14

RESULT 12
B0447972/c
LOCUS
DEFINITION UI-H-EUI-bah-k-20-0-UI.s1 NCI CGAP Ctl1 Homo sapiens cDNA clone
UI-H-EUI-bah-k-20-0-UI 3', mRNA sequence.
ACCESSION B0447972
VERSION B0447972.1 GI:21251084
KEYWORDS EST.

SOURCE
ORGANISM

Homo sapiens (human)
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 504)
NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

REFERENCE
AUTHORS
TITLE

JOURNAL
COMMENT

FEATURES
source

Location/Qualifiers
1..504
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-EUI-bah-k-20-0-UI"
/tissue_type="Osteoarthritic Cartilage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP Ctl1"
/note="Organ: Knee; Vector: p7773-Pac (Pharmacia) with a
modified polylinker; Site1: EcoR I; Site2: Not I;
NCI CGAP Ctl1 is a normalized cDNA library containing the
following tissue(s): Osteoarthritic Cartilage. The library
was constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into p7773-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
TGATCACGCT.
TAG_TISSUE=osteoarthritic cartilage
TAG_LIB=UI-H-EUI
TAG_SEQ=TGATCACGCT

ORIGIN

Query Match 99.5%; Score 321.4; DB 5; Length 504;
Best Local Similarity 99.7%; Pred. No. 1.4e-81;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db 336 CGGGGCCAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 277
Qy 61 CATCCTCATGACCATTTGATTCACCTCTTACTGTTACTGGTCTGTGACAGTCTACCA 120
Db 276 CATCCTCATGACCATTTGATTCACCTCTTACTGTTACTGGTCTGTGACAGTCTACCA 217
Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCTGCAA 180
Db 216 TATTATGTTACAGATAATACAGGAAAAACGGGGTTTACTAGTAGCGGCCCATAGCTGCAA 157
Qy 181 CTTTGGCACTCAGTGCATGTCGGCCCTGACGCTGGGGCTGTTGCCCTGCCCCCT 240
Db 156 CTTTGGCACTCAGTGCATGTCGGCCCTGACGCTGGGGCTGTTGCCCTGCCCCCT 97
Qy 241 TGGTCTGCCCTAGATACAGGATTTATACCCACACCTGTCTACAGTGTCAATCAAT 300
Db 96 TGGTCTGCCCTAGATACAGGATTTATACCCACACCTGTCTACAGTGTCAATCAAT 37

```

QY 301 AAAGTGCACGTGCTTGTGAAAAA 323
    |||||||
Db 36 AAAGTGCACGTGCTTGTGAAAAA 14

RESULT 13
BM672499/c
LOCUS
DEFINITION
    BM672499 505 bp mRNA linear EST 27-FEB-2002
    UI-E-CQ1-ada-d-12-0-UI.s1 UI-E-CQ1 Homo sapiens cDNA clone
    UI-E-CQ1-ada-d-12-0-UI 3', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
    BM672499.1 GI:18982397
    Homo sapiens (human)
ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Hominidae; Homo.
    1 (bases 1 to 505)
    Bernaldo,M.F., Lennon,G. and Soares,M.B.
    Normalization and subtraction: two approaches to facilitate gene
    discovery
JOURNAL
PUBMED
    Genome Res. 6 (9), 791-806 (1996)
COMMENT
    Contact: Soares, MB
    Coordinated Laboratory for Computational Genomics
    University of Iowa
    375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
    Tel: 319 335 8250
    Fax: 319 335 9565
    Email: bento-soares@uiowa.edu
    Tissue Procurement: Dr. Gregg Hageman
    cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
    cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
    DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
    Clones Distribution: Researchers may obtain clones from Research
    Genetics (www.resgen.com).
    Seq primer: M13 Forward
    POLYA=Yes.
FEATURES
    source
    1. 505
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="UI-E-CQ1-ada-d-12-0-UI"
        /dev_stage="adult"
        /tissue_type="optic nerve"
        /lab_hosts="DH10B (Life Technologies) (T1 phage resistant)"
        /clone_lib="UI-E-CQ1"
        /note="Organ: eye; Vector: pTT3-Pac (Pharmacia) with a
        modified polylinker; Site_1: EcoR I; Site_2: Not I;
        UI-E-CQ1 is a normalized cDNA library containing the
        following tissue(s): optic nerve. The library was
        constructed according to Bernaldo, Lennon and Soares,
        Genome Research, 6:791-806, 1996. First strand cDNA
        synthesis was primed with an oligo-dT primer containing a
        Not I site. Double stranded cDNA was ligated to an EcoR I
        adaptor, digested with Not I, and cloned directionally
        into pTT3-Pac vector. The oligonucleotide used to prime
        the synthesis of first-strand cDNA contains a library tag
        sequence that is located between the Not I site and the
        (dT)18 tail. The sequence tag for this library is
        CCATTAGTG. This library was created for the program, Gene
        Discovery in the Visual System, supported by National Eye
        Institute (NEI).
        TAG TISSUE=human optic nerve
        TAG LIB=UI-E-CQ1
        TAG_SEQ=CCATTAGTG"
ORIGIN
    Query Match 99.5%; Score 321.4; DB 3; Length 505;
    Best Local Similarity 99.7%; Pred. No. 1.4e-81;
    Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 CGGGCCCGAGGCTATGCTCCACCGCAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
    |||||||
Db 336 CGGGCCCGAGGCTATGCTCCACCGCAGTGCCTGAACATCTGGGCCCTGATTCGGG 277

QY 61 CATCCTCATGACCATTTGGATTACCCCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 120
    |||||||
Db 276 CATCCTCATGACCATTTGGATTACCTCTGTTACTGGTATTTCGGCTCTGTGACAGTCTACCA 217

QY 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCCATAGCTGCAA 180
    |||||||
Db 216 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCCGCCCATAGCTGCAA 157

QY 181 CTTTTCACCTCCACTGTGCAATCTGSCCTGCACCGCTGGGGTGTGGCCCTGCCCCCT 240
    |||||||
Db 156 CTTTTCACCTCCACTGTGCAATCTGSCCTGCACCGCTGGGGTGTGGCCCTGCCCCCT 97

QY 241 TGGTCTCTGCCCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 300
    |||||||
Db 96 TGGTCTCTGCCCCCTAGATACAGCAGTTTATATACCCACACACCTGTCTACAGTGTCAATCAAT 37

QY 301 AAAGTGCACGTGCTTGTGAAAAA 323
    |||||||
Db 36 AAAGTGCACGTGCTTGTGAAAAA 14

RESULT 14
BM6767351
LOCUS
DEFINITION
    K-EST0049808 S6SNU620s1 Homo sapiens cDNA clone S6SNU620s1-13-C12
    5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
    BM6767351.1 GI:19096966
    Homo sapiens (human)
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
    Hominidae; Homo.
    1 (bases 1 to 510)
    Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
    Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
    Kim,Y.S.
    21C Frontier Korean EST Project 2001
    Unpublished (2002)
    Contact: Kim YS
    Genome Research Center
    Korea Research Institute of Bioscience & Biotechnology
    52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
    Tel: +82-42-860-4470
    Fax: +82-42-860-4409
    Email: yongseung@mail.kribb.re.kr
    Plate: 13 row: C column: 12
    High quality sequence stop: 510.
FEATURES
    source
    1. 510
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="S6SNU620s1-13-C12"
        /sex="g"
        /tissue_type="Asclites"
        /cell_type="Scattering floating"
        /cell_line="SNU-620"
        /lab_hosts="Top10F"
        /clone_lib="S6SNU620s1"
        /note="Organ: Stomach; Vector: pcNS; Site 1: EcoRI;
        Site 2: NotI. The poly (A) + RNA was dephosphorylated with
        bacterial alkaline phosphatase (BAP) and then decapped
        with tabacco acid pyrophosphatase (TAP). The decapped
        intact mRNA was ligated with DNA-RNA linker including EcoR
        I site by treatment of T4 RNA ligase and the first strand
        cDNA was synthesized from oligo dT-selected mRNA by

```


priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 - 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and N(drt)14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F' with electroporation method."

ORIGIN

```

Query Match      99.5%; Score 321.4; DB 3; Length 510;
Best Local Similarity 99.7%; Pred. No. 1.4e-81;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTTCACGGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db 182 CGGGGCCAGGCTATGCTTCACGGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 241

Qy 61 CATCTCATGACCATGTGATTCACCTGTTACTGTATTTCGCTCTGAGCAGTCTACCA 120
Db 242 CATCTCATGACCATGTGATTCACCTGTTACTGTATTTCGCTCTGAGCAGTCTACCA 301

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGCCCATAGCTGCAA 180
Db 302 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGCCCATAGCTGCAA 361

Qy 181 CTTTGGCACTCAGTGTGAATGTGGGCCCTGACGCTGGGGCTGTGGCCCTGCCCT 240
Db 362 CTTTGGCACTCAGTGTGAATGTGGGCCCTGACGCTGGGGCTGTGGCCCTGCCCT 421

Qy 241 TGGTCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCTTCAAT 300
Db 422 TGGTCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCTTCAAT 481

Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323
Db 482 AAAGTGCACGTGCTTGTGAAAAA 504

```

RESULT 15

BU783190/c

LOCUS in04d02.x1 Human insulinoma Homo sapiens cDNA clone IMAGE:6123315
DEFINITION 3' similar to SW:IN19_HUMAN P13164 INTERFERON-INDUCIBLE PROTEIN
9-27 ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

EST. BU783190.1 GI:23827129
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 522)
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marr,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Teagareishvili,R.,

Williams,T., Jackson,Y. and Bowers,Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@bioh.harvard.edu

Library was constructed by Dr. J. J. Ferrer In vivo mass-excised to

phuescript SK- by Dr. H. Inoue DNA sequencing by: Washington

University Genome Sequencing Center For information on obtaining a

clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)

Seq primer: -40UP from Gibco

High quality sequence stop: 403.

FEATURES

source

```

1..522
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6123315"
/tissue_type="insulinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Human insulinoma"
/notes="Organ: pancreas; Vector: pBluescript SK-; Site 1:
XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
pBluescript SK- by Dr. H. Inoue following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Hiroshi Inoue, MD/PhD for further
information on this library (Metabolism Division, Permutt
Laboratory, Washington University School of Medicine, Box
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
is a Washington University Pancreas EST project library."

```

ORIGIN

```

Query Match      99.5%; Score 321.4; DB 5; Length 522;
Best Local Similarity 99.7%; Pred. No. 1.4e-81;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGGCCAGGCTATGCTTCACGGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 60
Db 334 CGGGGCCAGGCTATGCTTCACGGCAAGTGCCTGAACATCTGGGCCCTGATTCGGG 275

Qy 61 CATCTCATGACCATGTGATTCACCTGTTACTGTGTTATTCGGCTCTGTGACAGTCTACCA 120
Db 274 CATCTCATGACCATGTGATTCACCTGTTACTGTGTTATTCGGCTCTGTGACAGTCTACCA 215

Qy 121 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGCCCATAGCTGCAA 180
Db 214 TATTATGTTACAGATAATACAGGAAAAACGGGGTTACTAGTAGCGCCCATAGCTGCAA 155

Qy 181 CTTTGGCACTCAGTGTGAATGTGGGCCCTGACGCTGGGGCTGTGGCCCTGCCCT 240
Db 154 CTTTGGCACTCAGTGTGAATGTGGGCCCTGACGCTGGGGCTGTGGCCCTGCCCT 95

Qy 241 TGGTCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCTTCAAT 300
Db 94 TGGTCTGCCCTAGATACAGAGTTTATACCCACACACCTGTCTACAGTGTCTTCAAT 35

Qy 301 AAAGTGCACGTGCTTGTGAAAAA 323
Db 34 AAAGTGCACGTGCTTGTGAAAAA 12

```

Search completed: December 7, 2005, 09:55:54

Job time : 3689.9 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 2600.1 Seconds
(without alignments)
6427.437 Million cell updates/sec

Title: US-09-980-046B-7

Perfect score: 294

Sequence: 1 cggcgatcgccggcggttat.....agcgagcacctagacaaaaa 294

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5893141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_px.*

9: gb_ro.*

10: gb_scs.*

11: gb_sy.*

12: gb_un.*

13: gb_vi.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	270.8	92.1	1736	8 BC030541	BC030541 Homo sapi
2	270.8	92.1	2271	8 AB010427	AB010427 Homo sapi
3	270.8	92.1	4491	8 HS9808338	BX648190 Homo sapi
4	269.4	91.6	2369	6 AR064037	AR064037 Sequence
5	269.4	91.6	2741	6 CQ850514	CQ850514 Sequence
6	269.4	91.6	2741	8 AK127665	AK127665 Homo sapi
7	269.4	91.6	2743	8 AB209746	AB209746 Homo sapi
8	269.4	91.6	2825	8 BC000201	BC000201 Homo sapi
9	269.4	91.6	2922	6 CQ717617	CQ717617 Sequence
10	269.4	91.6	2951	8 BC002489	BC002489 Homo sapi
11	269.4	91.6	199450	8 AC005674	AC005674 Homo sapi
12	269.4	91.6	240000	14 AC009528	AC009528 Homo sapi
13	267.8	91.1	2850	8 AF020056	AF020056 Homo sapi
14	267.4	91.0	3075	10 BV179203	BV179203 sqm10429
15	256.4	87.2	2047	8 AB168915	AB168915 Macaca fa
16	181	61.6	2115	8 AK222764	AK222764 Homo sapi
17	159	54.1	2125	8 AK222897	AK222897 Homo sapi
18	152.6	51.9	212722	14 AC148123	AC148123 Ootlemur

c	19	134.8	45.9	207389	14	AC148146
	20	111.6	38.0	2856	9	BC049117
c	21	111.6	38.0	104872	9	AC084322
c	22	111.6	38.0	347166	9	AC084070
	23	110.8	37.7	2908	9	AF020055
c	24	108.8	37.0	177293	14	AC148073
	25	101.4	34.5	2844	9	BC085864
c	26	101.4	34.5	110000	14	AC091339_01
c	27	101.4	34.5	110000	14	AC091339_02
	28	93.4	31.8	2878	9	AB084171
c	29	80	27.2	165573	14	AC121673
c	30	80	27.2	174283	14	AC137259
c	31	80	27.2	189838	14	AC121037
	32	41.4	14.1	3297	5	AF020054
	33	38.2	13.0	253631	14	AC126155
c	34	37.4	12.7	160238	9	AC133502
c	35	37	12.6	224778	14	AC079491
	36	36.6	12.4	254878	14	AC110841
	37	35.8	12.2	124645	8	HUAC004638
	38	35.8	12.2	124998	8	AC026397
	39	35.8	12.2	128877	14	AC026399
	40	35.8	12.2	170691	8	AC044802
	41	35.8	12.2	206102	8	AC027131
c	42	35.4	12.0	12768	9	AF263346
	43	35.4	12.0	112467	14	AC164519
c	44	35	11.9	78034	8	AL358473
c	45	35	11.9	157351	9	AC156940

ALIGNMENTS

BC030541 1736 bp mRNA linear PRI 12-OCT-2004
Homo sapiens WD repeat domain 1, transcript variant 2, mRNA (cdna
clone MGC:40281 IMAGE:4403484), complete cds.

BC030541

BC030541.2 GI:34192109

MGC.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 1736)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,

Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,

Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,

Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,

Sheezy,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,

Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,

Abramson,R.D., Mullahy,S.J., Bosak,S.A., McSwan,P.J.,

McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,

Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,

Faney,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,

Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,

Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,

Butterfield,Y.S., Krzywicki,M.I., Skalska,U., Smalilov,D.E.,

Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1736)

Director MGC Project.

Direct Submission

Submitted (07-MAY-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

TITLE Full-length human cDNA
JOURNAL Patent: EP 1447413-A 983 18-AUG-2004;
Research Association for Biotechnology (JP)
FEATURES Location/Qualifiers
source 1. .2741
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 91.6%; Score 269.4; DB 6; Length 2741;
Best Local Similarity 99.6%; Pred. No. 1.1e-77;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 21 TCCATGATGCTCTGTCAAGGAGTGACAAATCACTACTGAGGAGCCACCCCGCCT 80
DB 1803 TCCATGATGCTCTGTCAAGGAGTGACAAATCACTACTGAGGAGCCACCCCGCCT 1862
QY 81 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGACGCGGAACATGTCTATT 140
DB 1863 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGACGCGGAACATGTCTATT 1922
QY 141 TCTGTGACGCGCCCATGCCCCACCCACCAAGAGCAGGAGGCCAGTCATGAC 200
DB 1923 TCTGTGACGCGCCCATGCCCCACCCACCAAGAGCAGGAGGCCAGTCATGAC 1982
QY 201 CCTCGTCTCTGCAGGGTGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTTC 260
DB 1983 CCTCGTCTCTGCAGGGTGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTTC 2042
QY 261 CACATGAAAAATAAAGCGAGCACCCTAGACAA 291
DB 2043 CACATGAAAAATAAAGCGAGCACCCTAAACAA 2073
RESULT 6
AK127665
LOCUS Homo sapiens cDNA FLJ45763 fis, clone N1ESE2000698, highly similar
DEFINITION to WD-repeat protein 1.
ACCESSION AK127665
VERSION 1 GI:34534673
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K.,
Arita, M., Mueashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Wakamatsu, A., Umezawa, A., Imabayashi, H.,
Fukuma, M., Izawa, Y., Hata, J., Ishii, S., Yamamoto, J., Isono, Y.,
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatauma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B.,
Suzuki, Y., Sugano, S., Nagahara, K., Masuno, Y., Nagai, K. and
Icogai, T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2741)
Icogai, T. and Yamamoto, J.
AUTHORS Direct Submission
TITLE Submitted (15-JUL-2003) Takao Iicogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomic@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and

RAB; annotation: Reverse Proteomics Research Institute, HRI and
RAB.
FEATURES Location/Qualifiers
source 1. .2741
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="N1ESE2000698"
/cell_type="mesenchymal stem cells"
/clone_lib="N1ESE2"
/notes="cloning vector: pME18SFL3"
ORIGIN
Query Match 91.6%; Score 269.4; DB 8; Length 2741;
Best Local Similarity 99.6%; Pred. No. 1.1e-77;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 21 TCCATGATGCTCTGTCAAGGAGTGACAAATCACTACTGAGGAGCCACCCCGCCT 80
DB 1803 TCCATGATGCTCTGTCAAGGAGTGACAAATCACTACTGAGGAGCCACCCCGCCT 1862
QY 81 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGACGCGGAACATGTCTATT 140
DB 1863 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGACGCGGAACATGTCTATT 1922
QY 141 TCTGTGACGCGCCCATGCCCCACCCACCAAGAGCAGGAGGCCAGTCATGAC 200
DB 1923 TCTGTGACGCGCCCATGCCCCACCCACCAAGAGCAGGAGGCCAGTCATGAC 1982
QY 201 CCTCGTCTCTGCAGGGTGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTTC 260
DB 1983 CCTCGTCTCTGCAGGGTGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTTC 2042
QY 261 CACATGAAAAATAAAGCGAGCACCCTAGACAA 291
DB 2043 CACATGAAAAATAAAGCGAGCACCCTAAACAA 2073
RESULT 7
AB209746
LOCUS Homo sapiens mRNA for WD repeat-containing protein 1 isoform 1
DEFINITION variant protein.
ACCESSION AB209746
VERSION 1 GI:62089077
KEYWORDS FLJ CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
Totoki, Y., Toyoda, A., Takeda, T., Sakaki, Y., Tanaka, A., Yokoyama, S.,
Ohara, O., Nagase, T. and F. Kikuno, R.
AUTHORS None Title
TITLE Published Only in Database (2005)
JOURNAL 2 (bases 1 to 2743)
AUTHORS Totoki, Y., Toyoda, A., Takeda, T., Sakaki, Y., Tanaka, A., Yokoyama, S.,
Ohara, O., Nagase, T. and F. Kikuno, R.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2005) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
Chiba, 292-0818, Japan (E-mail: cdnainfo@kazusa.or.jp,
URL: http://protein.gsc.riken.go.jp/, Tel: 81-438-52-3930,
Fax: 81-438-52-3931)
COMMENT This work was supported in part by the National Project on Protein
Structural and Functional Analysis, Ministry of Education, Culture
Sports, Science and Technology of Japan. Totoki Y., Toyoda A.,
Takeda T., Sakaki Y., Tanaka A., Yokoyama S. RIKEN Genomic Sciences
Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan.
e-mail: aktanaka@postman.riken.go.jp
URL: http://protein.gsc.riken.go.jp/.
FEATURES Location/Qualifiers

source	1. .2743 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="ek00180" /tissue_type="aorta endothelial cell" /note="This clone is also named as hek003002109 vector:pBlueScriptII SK plus" <1. .2743 /gene="WD repeat-containing protein 1 isoform 1 variant" <1. .1977 /gene="WD repeat-containing protein 1 isoform 1 variant" /note="Start codon is not identified." /codon_start=3 /evidence=not experimental /product="WD Repeat-containing protein 1 isoform 1 variant" /protein_id="BAD92983.1" /db_xref="GI:62089078" /translation="RAGPGSGRRARRGNKWMRPYEIKKVPASLPQVERGVSKIIGD PKGNFLVTNGKCVILNRINDPALADIVTEHAHQVWVAKYAPSGFYIAGDVSGLKRI WDTTQKEHLLEYQPPAGKIKDIANTEDSKRIAVGEGREKFGAVFLWDSVSGEI TGNKRVINSVDIKSGRPLATGSDNCAFFEGPPKFTKFTGDRFRVNCVRFSD GNRFATASADGOIYIYDGKTGEKVGALGSKAHDGGIYAISSPDSTHLLSASDKTS KINDVSNVSVSTPFPMGSTDVLQGLQKQWHLLSVLSGYINLYDRNPFSGELHWI KGHSKISICLTVHNGKSYIYSGSDGHIYNWDSYTGENDSPAGKGTNVCISQIIVLL ESGOLISCSMDTVRYTSLMRDYSGQGVKLDVQPKCVAVGPGGYAVVVCIGIIVLL KDQRKPSIDNPGEVPEVAVHGGDTVAIGVDPGNVRLYSILCTTLKDEGKLEANG PVTDAVSHDGAFLVCDASKVTVTVFSVADGTSENNVFGHHKIVCLAWSFDNEHFA SGGMDMMVYVWTLSDPETRVKIQDAHLRHVSSLAWLDEHTLVTTSHDASVKWETITY "
gene	
CDS	
FEATURES	source 1. .2825 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:3118 IMAGE:3351638" /tissue_type="Eye, retinoblastoma" /clone_lib="NIH MGC_16" /lab_host="DH10B-R" /note="Vector: pOTB7" 1. .2825 /gene="WDRI" /note="synonym: NORI-1" /db_xref="GeneID:9948" /db_xref="MIM:604734" 83_1903 /gene="WDRI" /codon_start=1 /product="WD repeat-containing protein 1, isoform 1" /protein_id="AAH00201.1" /db_xref="GI:12652891"
ORIGIN	
Query Match	91.6%; Score 269.4; DB 8; Length 2743;
Best Local Similarity	99.6%; Pred. No. 1.1e-77;
Matches 270; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Qy	21 TCCCATGATGCTCTGTCAAGAGTGGACAATCACCTACTGAGGAGCCCCACCCCGCCT 80
Db	 1836 TCCCATGATGCTCTGTCAAGAGTGGACAATCACCTACTGAGGAGCCCCACCCCGCCT 1895
Qy	81 CTGGATGACCGGAATCAGGACTAGAGTTTAACTGCAGCGGAACATGTCATTTCTCTATT 140
Db	1896 CTGGATGACCGGAATCAGGACTAGAGTTTAACTGCAGCGGAACATGTCATTTCTCTATT 1955
Qy	141 TCTGTGACGCGCCCCCATGCCCCACCCACCAAGAGGAGGAGGCGGCCCATGTCATC 200
Db	1956 TCTGTGACGCGCCCCCATGCCCCACCCACCAAGAGGAGGAGGCGGCCCATGTCATC 2015
Qy	201 CCTGCTCTCTGAGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGATTG 260
Db	2016 CCTGCTCTCTGAGGGTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGATTG 2075
Qy	261 CACATGAAAAATAAAGCGAGCACCTAGACAA 291
Db	 2076 CACATGAAAAATAAAGCGAGCACCTAAACAA 2106
RESULT 8	
BC000201	
LOCUS	BC000201 2825 bp mRNA linear PRI 29-JUN-2004
DEFINITION	Homo sapiens WD repeat domain 1, transcript variant 1, mRNA (cDNA clone MGC:3118 IMAGE:3351638), complete cds.
ACCESSION	BC000201
VERSION	BC000201.1 GI:12652890
KEYWORDS	MGC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 2825)
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

/db_xref="GeneID:9948"
/db_xref="WIM:604734"
/translation="MPVEIKKVFASLPQVERGVSKLIGDPKGNFLYTNKCVILRN
IDNPALADIYTEHAHQVAVKAPSPFYIASGVSKLRINDWTIQEHLKLYQBPFA
GKIKDIAMTSDSKIAVVGGRKFGAVFLWDSGSSVGEITGHNKVINSDIHKQSRPY
RLATGSDNCAAFEGEPFKFTVGDHSFRVFPSPDGNRFATASADGQIYIYDG
KTGBKCALGSKAHADGIIYAIWSPPDSTHLLSASGDKTSKIDWSVNSVVSFPMGS
TVLDQGLCLWOKDHLLSVLSGVINYLDNRNPSKPLHVIKHSKSIQCLIVHKNKGK
SYLISGSHDGHINWSETGENDSFAGKHTNQVSRMTVDSEQLISCSDMDTVRVS
LMURDSGGQVKLDQVFKCAVPGPGIAVVCIGIQLVKDQKCFSDINPGYEPFV
VAVHGGDVAIGVDGNRLYLSILGTTLDEKGLLEAKGPFVDVAYSHDGAFLVCD
ASKVTVFVSADGVSENNFYGHAKIVCLAWSPDNEHFASGMDMMVYVWTLSDPET
RVKIQDAHRLHHVSSLAWLDEHTLVTTSHDSVKEWTITY"

ORIGIN

Query Match 91.6%; Score 269.4; DB 8; Length 2825;
Best Local Similarity 99.6%; Pred. No. 1.1e-77;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	21	TCCCATGATGCTCTGTCAAGGAGTGACATCACTACTGTAGGAGCCGCCACCCCGCCT	80
Db	1862	TCCCATGATGCTCTGTCAAGGAGTGACATCACTACTGTAGGAGCCGCCACCCCGCCT	1921
Qy	81	CTGGATGACCGAATCAGGAGTAGAGTTTAACTGACGCGAAACATGTCATTTCTCTATT	140
Db	1922	CTGGATGACCGAATCAGGAGTAGAGTTTAACTGACGCGAATGTCATTTCTCTATT	1981
Qy	141	TCTGTGACGGCGCCCATGCCCCACCCACCAAGAGGAGGAGGCGCCAGTCATGAC	200
Db	1982	TCTGTGACGGCGCCCATGCCCCACCCACCAAGAGGAGGAGGCGCCAGTCATGAC	2041
Qy	201	CCTCGTCTCTGCAGGGTGTGTGACAGTCTTCTGAAAGCTTTAGACAGTAACAGTTTG	260
Db	2042	CCTCGTCTCTGCAGGGTGTGTGACAGTCTTCTGAAAGCTTTAGACAGTAACAGTTTG	2101
Qy	261	CACATGAAAAATAAGCGAGCACCTAGACAA	291
Db	2102	CACATGAAAAATAAGCGAGCACCTAACA	2132

RESULT 9
CQ717617
LOCUS CQ717617 2922 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 3551 from Patent WO02068579.
ACCESSION CQ717617
VERSION CQ717617.1 GI:42278474
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 3551 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1..2922
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 91.6%; Score 269.4; DB 6; Length 2922;
Best Local Similarity 99.6%; Pred. No. 1.1e-77;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	21	TCCCATGATGCTCTGTCAAGGAGTGACATCACTACTGTAGGAGCCGCCACCCCGCCT	80
Db	1983	TCCCATGATGCTCTGTCAAGGAGTGACATCACTACTGTAGGAGCCGCCACCCCGCCT	2042

Qy	81	CTGGATGACCGAATCAGGAGTAGAGTTTAACTGACGCGAAACATGTCATTTCTCTATT	140
Db	2043	CTGGATGACCGAATCAGGAGTAGAGTTTAACTGACGCGAAACATGTCATTTCTCTATT	2102
Qy	141	TCTGTGACGGCGCCCATGCCCCACCCACCAAGAGGAGGAGGCGCCAGTCATGAC	200
Db	2103	TCTGTGACGGCGCCCATGCCCCACCCACCAAGAGGAGGAGGCGCCAGTCATGAC	2162
Qy	201	CCTCGTCTCTGCAGGGTGTGTGACAGTCTTCTGAAAGCTTTAGACAGTAACAGTTTG	260
Db	2163	CCTCGTCTCTGCAGGGTGTGTGACAGTCTTCTGAAAGCTTTAGACAGTAACAGTTTG	2222
Qy	261	CACATGAAAAATAAGCGAGCACCTAGACAA	291
Db	2223	CACATGAAAAATAAGCGAGCACCTAACA	2253

RESULT 10

LOCUS BC002489
DEFINITION Homo sapiens WD repeat domain 1, transcript variant 1, mRNA (CDNA
clone MGC:1382 IMAGE:3051319), complete cds.
ACCESSION BC002489
VERSION BC002489.2 GI:33876752
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

AUTHORS 1 (bases 1 to 2951)
Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD,
Collins FS, Wagner L, Shenmen CW, Schuler GD, Altschul SF, Zeeberg
B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T,
Max SI, Wang J, Haieih F, Diatchenko L, Marusina K, Farmer AA, Rubin
GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL,
Scheetz TE, Brownstein MJ, Uddin TB, Toshiyuki S, Carninci P,
Prange C, Raha SA, Loquellano NA, Peters GJ, Abramson RD, Mullahy
SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH,
Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Rulyk SW,
Villalón DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J,
Helton E, Kettman M, Madan A, Rodrigues S, Sanchez A, Whiting M,
Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakealey RW,
Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, U,
Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska K, U,
Smalilus DE, Schnerch A, Schein JE, Jones SJ, Jones SJ and Marra MA.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932

CONSRTM

TITLE human and mouse cDNA sequences
JOURNAL
PUBMED 12477932
REFERENCE 2 (bases 1 to 2951)
AUTHORS Director MGC Project.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT On Aug 19, 2003 this sequence version replaced gi:12803340.
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (ILLNL)
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,

QY	261	CACATGAAATAAAGCGGACCTAGACAA	291
DB	2131	CACATGAAATAAAGCGGACCTAAACAA	2161
RESULT 14			
BV179203/c			
LOCUS			
DEFINITION			
ACCESSION	BV179203	3075 bp	DNA linear STS 10-JUN-2004
VERSION	BV179203.1	GI:48015861	sqmm104292 Human DNA (Sequenom) Homo sapiens STS genomic, sequence tagged site.
KEYWORDS	STS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M., Cantor,C.R. and Braun,A.		
TITLE	Large-Scale Validation of Single Nucleotide Polymorphisms in Gene Regions		
JOURNAL	Genome Res. (2004) In press		
COMMENT	Contact: Andreas Braun Pharmaceuticals division Seqenom, Inc. 3595 John Hopkins Court, San Diego, CA 92121, USA Tel: 18582029018 Fax: 18582029020 Email: abraun@sequenom.com Primer A: No primer sequence submitted Primer B: No primer sequence submitted STS size: 3075		
FEATURES			
source	Location/Qualifiers 1..3075 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /clone_lib="Human DNA (Sequenom)" <1..>3075		
STS ORIGIN			
Query Match	91.0%;	Score 267.4;	DB 10; Length 3075;
Best Local Similarity	98.9%;	Pred. No. 4.9e-77;	
Matches 288;	Conservative	0; Mismatches 3;	Indels 0; Gaps 0;
QY	21	TCCCATGATGCCTCTGTCAAGGAGTGGAACAATCACCTACTGAGGAGCCCCACCCC	80
DB	1082	TCCCATGATGCCTCTGTCAAGGAGTGGAACAATCACCTACTGAGGAGCCCCACCCC	1023
QY	81	CTGGATGGACCGAATCAGGGACTAGAGTTTAATCTGACGGGAACATGTCTATTCTCTATT	140
DB	1022	CTGGATGGACCGAATCAGGGACTAGAGTTTAACTGACGGGAACATGTCTATTCTCTATT	963
QY	141	TCGTGTACGCCGCCCCCATGCCACCACCAACAGAGAGCGAGGCGCCAGTCATGAC	200
DB	962	TCGTGTACGCCGCCCCCATGCCACCACCAACAGAGAGCGAGGCGCCAGTCATGAC	903
QY	201	CCTCGTCTCTGCAGGTGTCTGTACACGTTCTTCTGAAAAGCTTTTAGACAGTAAACAGTTTG	260
DB	902	CCTNGTCTCTGCAGGTGTCTGTACACGTTCTTCTGAAAAGCTTTTAGACAGTAAACAGTTTG	843
QY	261	CACATGAAAAATAAAGCGGACCTAGACAA	291
DB	842	CACATGAAAAATAAAGCGGACCTAAACAA	812
RESULT 15			
AB168915			
LOCUS			
DEFINITION			
AB168915	2047 bp	mRNA linear	PRI 18-JUN-2005
Macaca fascicularis testis cDNA, clone: QtsA-15761, similar to			

human WD repeat domain 1 (WDRL1), transcript variant 1, mRNA,
RefSeq: NM_017491.2.
AB168915
AB168915.1 GI:67969327
oligo capping: fis (full insert sequence).
Macaca fascicularis (crab-eating macaque)
Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecinae; Macaca.

1 International consortium for macaque cDNA sequencing and analysis:
DNA sequences of macaque genes expressed in brain or testis and its
evolutionary implications
Unpublished

2 Oeada, N., Hirata, M., Tanuma, R., Kusuda, J., Hida, M., Suzuki, Y.,
Sugano, S., Gojobori, T., Shen, J.C.-K., Wu, C. I. and Hashimoto, K.
Substitution rate and structural divergence of 5'UTR evolution:
Comparative analysis between human and cynomolgus monkey cDNAs
Unpublished
3 (bases 1 to 2047)
Hashimoto, K., Kusuda, J. and Sugano, S.

Custom primers were used for 5' and 3'-end sequencing. The full-insert sequencing was done by primer-walking method using ABI DNA sequencer.

FEATURES	source	Location/Qualifiers
		1..2047
		/organism="Macaca fascicularis"
		/mol_type="mRNA"
		/db_xref="taxon:9541"
		/clone="QtsA-15761"
		/sex="male"
		/clone_lib="macaque cDNA library QtsA"
		/dev_stage="adult"
CDS		730..1731

/notes="unnamed protein product; Homo sapiens WD repeat domain 1 (WDR1), transcript variant 1, mRNA, RefSeq: NM_017491.2"	
/codon_start=1	
/protein_id="BAE01016.1"	
/db_xref="GI:67969328"	
/translation="MGSTVLDQQLGCLMQKHLLSVLSGYINLYDRNPSPKPLRVIKGHSKIQCIITVHNKGKSYISGSHDGHYNWDSGENDSPAGKGTHTQVSRMTVDESGQLISCSMDITVRYTSLRDYSGQGVVLLDQPKCVAPGPGYAVVVCIGQIVLLKDGKRCPSIDNPGEYVPEVAVHPGDDTVAGGADVGNVLYSLGTLTKDEGKLLEAKPGVTDVAVSHDGAFLAVCDASKVTVFVSADGYSNNVFGYGHAKIICLAWSPDNHEFASGGMDMMVYVWTLSDPETRVKIQDAHLRHHVSSLAWLDEHTLVTSHDASVKWEHTIY"	
Query Match	87.2%; Score 256.4; DB 8; Length 2047;
Best Local Similarity	96.0%; Pred. No. 2.1e-73;
Matches 263; Conservative	0; Mismatches 11; Indels 0; Gaps 0;
Qy	21 TCCCATGATGCCTCTGTCAAGGAGTGGACAAATCACCTACTGTAGGAGGCCCCACCCCGCCT 80
Db	1690 TCCCATGACGCCTCTGTCAAGGAGTGGACAAATCACCTACTGTAGGAGGCCCCACCCCGCCT 1749
Qy	81 CTGGATGGACCGCAATCAGGGACTAGAGTTTAACTGCAGCGGAACATGTCTATTCTCTATT 140
Db	1750 CTGGATGGACCGCAATCAGGGACTAGCGTTTAACTGCAGCGGAACATGTCTTCTCTATT 1809
Qy	141 TCTGTGACGCGCCCCCATGCCCCACCCACCAACAAGAGGACAGAGGCCCCAGTCATGAC 200
Db	1810 TCTGTGACGCGCCCCCATGCCCCCGCCCACTGCAAGAGGACAGAGGCCCCCGTCATGAC 1869
Qy	201 CCTCGTCTCTGCAGGTGTCTGTACAGGTCCTTCTGAAAGCTTTAGACAGTACACGTTTG 260
Db	1870 CCTCGTCTCTGCAGGTGTCTGTACAGGTCCTTCTGAAAGCTTTAGACAGGAAACAGTTTG 1929
Qy	261 CACATGAAAATAAAGCGGACACCTAGACAAAAA 294
Db	1930 CACATGAAAATAAAGCAGCACCTAAAAAANA 1963

Search completed: December 7, 2005, 04:22:08

Job time : 2604.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:31:40 ; Search time 732.068 Seconds
(without alignments)
3320.997 Million cell updates/sec

Title: US-09-980-046B-7
Perfect score: 294
Sequence: 1 cggcgatgcggcggttat.....agcgagcactagacaaaaa 294

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	270.8	92.1	2271	9	US-10-756-149-1200
2	269.4	91.6	3047	9	US-10-450-763-16835
3	267.8	91.1	603	5	US-10-060-036-44
4	253.6	86.3	452	3	US-09-918-995-35013
5	244.4	83.1	485	3	US-09-918-995-17083
6	66.8	22.7	5800	6	US-10-062-674-2202
7	63.2	21.5	506	7	US-10-191-803-483
8	34.2	11.6	598	5	US-10-027-632-321061
9	34.2	11.6	598	6	US-10-027-632-321061
10	32.8	11.2	941	5	US-10-123-155-464
11	32.8	11.2	941	6	US-10-146-731-464
12	32.8	11.2	941	6	US-10-140-472-464
13	32.8	11.2	941	6	US-10-141-761-464
14	32.8	11.2	941	6	US-10-142-885-464
15	32.8	11.2	941	6	US-10-158-790-464
16	32.8	11.2	941	6	US-10-137-871-464
17	32.8	11.2	941	6	US-10-140-923-464
18	32.8	11.2	941	6	US-10-141-756-464
19	32.8	11.2	941	6	US-10-141-759-464
20	32.8	11.2	941	6	US-10-140-805-464
21	32.8	11.2	941	6	US-10-140-864-464
22	32.4	11.0	1559	8	US-10-787-284-2
23	32.4	11.0	2724	3	US-09-938-842A-344

c	24	32.4	11.0	2724	3	US-09-938-842A-344	Sequence 344, App
c	25	32.2	11.0	1143	7	US-10-437-963-87137	Sequence 87137, A
c	26	32.2	11.0	5589	7	US-10-437-963-56243	Sequence 56243, A
c	27	32.2	11.0	120239	7	US-10-322-281-271	Sequence 271, App
c	28	32	10.9	533	5	US-10-027-632-191275	Sequence 191275,
c	29	32	10.9	533	6	US-10-027-632-191275	Sequence 191275,
c	30	32	10.9	551	5	US-10-027-632-289809	Sequence 289809,
c	31	32	10.9	551	6	US-10-027-632-289809	Sequence 289809,
c	32	31.6	10.7	708	5	US-10-184-644-584	Sequence 584, App
c	33	31.6	10.7	708	5	US-10-184-644-584	Sequence 584, App
c	34	31.6	10.7	105219	5	US-10-087-192-658	Sequence 658, App
c	35	31.4	10.7	602	8	US-10-723-860-1882	Sequence 1882, Ap
c	36	31.4	10.7	602	9	US-10-756-149-1788	Sequence 1788, A
c	37	31.4	10.7	2000	7	US-10-260-238-1696	Sequence 1696, Ap
c	38	31.4	10.7	3931	3	US-09-764-846-304	Sequence 304, App
c	39	31.4	10.7	3931	5	US-10-091-483-304	Sequence 304, App
c	40	31.4	10.7	350764	5	US-10-087-192-1864	Sequence 1864, Ap
c	41	31.2	10.6	1400	9	US-10-956-157-5341	Sequence 5341, Ap
c	42	31.2	10.6	1457	9	US-10-956-157-106	Sequence 106, App
c	43	30.8	10.5	567	9	US-10-450-763-128	Sequence 128, App
c	44	30.8	10.5	777	10	US-11-090-847-213	Sequence 213, App
c	45	30.8	10.5	2111	8	US-10-357-930-22054	Sequence 22054, A

ALIGNMENTS

RESULT 1
US-10-756-149-1200
; Sequence 1200, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10756.149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1200
; LENGTH: 2271
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-10-756-149-1200

Query Match		92.1%	Score	270.8;	DB	9;	Length	2271;
Best Local Similarity		99.3%	Pred. No.	9.9e-87;				
Matches		272;	Conservative	0;	Mismatches	2;	Indels	0;
		Gaps						
Qy	21	TCCCATGATGCTCTGTCAAGGAGTGGAACAATCACCTACTGAGGAGCCCCACCCCGGCT	80					
Db	1982	TCCCATGATGCTCTGTCAAGGAGTGGAACAATCACCTACTGAGGAGCCCCACCCCGGCT	2041					
Qy	81	CTGGATGACCGAATCAGGAGCTAGAGTTTAACTGACGCGGAACATGTCATTCTCTATT	140					
Db	2042	CTGGATGACCGAATCAGGAGCTAGAGTTTAACTGACGCGGAACATGTCATTCTCTATT	2101					
Qy	141	TCTGTGACGCGCCCGCCATGCCCCACCCACACAGAGGAGGAGGCCCGCATGATGAC	200					
Db	2102	TCTGTGACGCGCCCGCCATGCCCCACCCACACAGAGGAGGAGGCCCGCATGATGAC	2161					
Qy	201	CCTCGTCTCTGACGGGTCTGTACAGCTTCTTCTTGAAGCTTTAGACAGTAACAGTTTG	260					
Db	2162	CCTCGTCTCTGACGGGTCTGTACAGCTTCTTCTTGAAGCTTTAGACAGTAACAGTTTG	2221					
Qy	261	CACATGAAAAATAAGCGAGCACCTAGACAAAAA	294					
Db	2222	CACATGAAAAATAAGCGAGCACCTAGACAAAAA	2255					

RESULT 2

US-10-450-763-16835
; Sequence 16835, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 16835
; LENGTH: 3047
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1530)..(2030)
; OTHER INFORMATION: 76% homologous to Homo sapiens 26-JUN-1997 cDNA encoding
; OTHER INFORMATION: human WD-40 protein, WDPro2, accession number Z92428_cdl, Smith-
; OTHER INFORMATION: Waterman Score=635.
US-10-450-763-16835

Query Match 91.6%; Score 269.4; DB 9; Length 3047;
Best Local Similarity 99.6%; Pred. No. 3.6e-86;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 21 TCCCATGATGCTCTGTCAAGGAGTGGAACAATCACTTACTTGAGGAGCCCCACCCCGCCT 80
Db 1989 TCCCATGATGCTCTGTCAAGGAGTGGAACAATCACTTACTTGAGGAGCCCCACCCCGCCT 2048

Qy 81 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 140
Db 2049 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 2108

Qy 141 TCTGTGACGGGCCCCCATGCCCCCACCACCAAGAGGAGGAGGCCCCAGTCATGAC 200
Db 2109 TCTGTGACGGGCCCCCATGCCCCCACCACCAAGAGGAGGAGGCCCCAGTCATGAC 2168

Qy 201 CCTGCTCTGCGAGGGTGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db 2169 CCTGCTCTGCGAGGGTGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 2228

Qy 261 CACATGAAAAATAAAGCGAGCACCCTAGACAA 291
Db 2229 CACATGAAAAATAAAGCGAGCACCCTAAACAA 2259

RESULT 3
US-10-060-036-44
; Sequence 44, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44

; LENGTH: 603
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-036-44

Query Match 91.1%; Score 267.8; DB 5; Length 603;
Best Local Similarity 99.3%; Pred. No. 7e-86;
Matches 269; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 21 TCCCATGATGCTCTGTCAAGGAGTGGAACAATCACTTACTTGAGGAGCCCCACCCCGCCT 80
Db 71 TCCCATGATGCTCTGTCAAGGAGTGGAACAATCACTTACTTGAGGAGCCCCACCCCGCCT 130

Qy 81 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 140
Db 131 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 190

Qy 141 TCTGTGACGGGCCCCCATGCCCCCACCACCAAGAGGAGGAGGCCCCAGTCATGAC 200
Db 191 TCTGTGACGGGCCCCCATGCCCCCACCACCAAGAGGAGGAGGCCCCAGTCATGAC 250

Qy 201 CCTGCTCTGCGAGGGTGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db 251 CCTGCTCTGCGAGGGTGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 310

Qy 261 CACATGAAAAATAAAGCGAGCACCCTAGACAA 291
Db 311 CACATGAAAAATAAAGCGAGCACCCTAAACAA 341

RESULT 4
US-09-918-995-35013
; Sequence 35013, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35013
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-35013

Query Match 86.3%; Score 253.6; DB 3; Length 452;
Best Local Similarity 98.2%; Pred. No. 8.6e-81;
Matches 267; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 21 TCCCATGATGCTCTGTCAAGGAGTGGAACAATCACTTACTTGAGGAGCCCCACCCCGCCT 80
Db 129 TCCCATGATGCTCTGTCAAGGAGTGGAACAATCACTTACTTGAGGAGCCCCACCCCGCCT 188

Qy 81 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 140
Db 189 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGCGAGCGGAACATGTCTCTATT 248

Qy 141 TCTGTGACGGGCCCCCATGCCCCCACCACCAAGAGGAGGAGGCCCCAGTCATGAC 200
Db 249 TCTGTGACGGGCCCCCATGCCCCCACCACCAAGAGGAGGAGGCCCCAGTCATGAC 308

Qy 201 CCTGCTCTGCGAGGGTGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTAACAGTTT 259
Db 309 CCTGCTCTGCGAGGGTGTCTGTACAGCTTCTTCTGAAAGCTTTAGACAGTAACAGTTT 368

Qy 260 GCACATGAAAAATAAAGCGAGCACCCTAGACAA 291
Db 311 GCACATGAAAAATAAAGCGAGCACCCTAAACAA 341

```
Db 369 GCACATGAAAAATAAAGCGACGACCTAAACAA 400
; OTHER INFORMATION: Incyte ID No. US20040005559A1 902527.2
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ... (5800)
; OTHER INFORMATION: a, t, c, g, or other
US-10-062-674-2202

Query Match 22.7%; Score 66.8; DB 6; Length 5800;
Best Local Similarity 77.4%; Pred. No. 1.1e-12;
Matches 212; Conservative 0; Mismatches 42; Indels 20; Gaps 10;

Qy 39 AAGGAGTGGACAATCACCT--ACTGAGGAGGCCCCACCCCGCCT-CTGGATGGACC--GA 93
Db 2253 AGGAGTGGACAATCACCTTACTGAGTGAGGCCCCACCCCGCCTACTGATGGACCTGAA 2312
Qy 94 ATCAGGAGCTAGAGTTTAACTGCA--GCGGAACATGTCTTCTCTATT--TCTGTGAGCG 150
Db 2313 TCAGGAGCTAGAGTTTAACTGCA--GCGGGAACATGTCTTCTCTATT--TCTGTGAGCG 2372
Qy 151 GCCCCATGCCCCACCCACCAACAGAG-----GCAGGAGGCCCCAG--TCATGACCCCTC 204
Db 2373 GCGCCCCAATGCCGCCACCCACCAACAGAGGCGCAGGAGGCCCCAGTTTCATGACCCCTC 2432
Qy 205 GTCTCT--GCAGGGTGTCTGTACAGGTTCTTCT--GAAAGCTTTTAGACAGTAACAGTTTGC 261
Db 2433 GTCTCTGCGCAGGGGTGTCTGTACAGGTTCTTCTTCTTGAAGGCTTTAGCAGAGTAACAGTTGTC 2492
Qy 262 ACATGA---AAAAATAAAGCGAGCACCTAGACAA 292
Db 2493 ACAGTAGAGGAATAAAGCGAGCACCTAAACAA 2526

RESULT 7
US-10-191-803-483/c
; Sequence 483 Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5090US
; CURRENT APPLICATION NUMBER: US/10/191,803
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 1140
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 483
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040014040A1 A1009012
US-10-191-803-483

Query Match 21.5%; Score 63.2; DB 7; Length 506;
Best Local Similarity 68.8%; Pred. No. 7.8e-12;
Matches 121; Conservative 0; Mismatches 43; Indels 12; Gaps 2;

Qy 121 GAACATGTCATTTCTCTATTCTGTGAGCGGCCCCCATGCCCCCACCACCCACACAGAGG 180
Db 504 GCACCTTCTTTTAACTATTCTGTAGCGCTCC-----CCTTCTGCTGCTGCTGGAG 451

Db 369 GCACATGAAAAATAAAGCGACGACCTAAACAA 400
; Sequence 17083, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17083
; LENGTH: 485
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ... (485)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-17083

Query Match 83.1%; Score 244.4; DB 3; Length 485;
Best Local Similarity 99.6%; Pred. No. 1.9e-77;
Matches 245; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 46 GGCAATCACCTACTAGGAGGCCCCACCCCGCCTCTGGATGGACCGGAATCAGGGACTAG 105
Db 75 GGCAATCACCTACTAGGAGGCCCCACCCCGCCTCTGGATGGACCGGAATCAGGGACTAG 134
Qy 106 AGTTAACTGACGCGGAACATGTCTATTCTTCTTCTGACGCGCCCCCATGCCCCCA 165
Db 135 AGTTAACTGACGCGGAACATGTCTATTCTTCTGACGCGCCCCCATGCCCCCA 194
Qy 166 CCCCACCAAGAGGCGGCCCCAGTCATGACCTCGTCTCTGCGAGGGTGTCTGTAC 225
Db 195 CCCCACCAAGAGGCGGCCCCAGTCATGACCTCGTCTCTGCGAGGGTGTCTGTAC 254
Qy 226 ACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTGCACATGAAAAATAAAGCGACCACT 285
Db 255 ACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTGCACATGAAAAATAAAGCGACCACT 314
Qy 286 AGACAA 291
Db 315 AAACAA 320

RESULT 6
US-10-062-674-2202
; Sequence 2202, Application US/10062674
; Publication No. US20040005559A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/062,674
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: US 09/625,102
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 2217
; SOFTWARE: PERL Program
; SEQ ID NO 2202
; LENGTH: 5800
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```



```
; SEQ ID NO 464
; LENGTH: 941
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-158-790-464

Query Match      11.2%; Score 32.8; DB 6; Length 941;
Best Local Similarity 9.3%; Pred. NO. 1;
Matches 24; Conservative 89; Mismatches 145; Indels 0; Gaps 0;

Qy 21 TCCCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTGAGGAGCCCCACCCCGCCT 80
Db HCBM...DABSSB..K.MT.C.N.W.HNWW.KAB.NC.W.DT..DTBABA.C.MSSTSS 339
      :|:: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|::: :|:::
Qy 81 CTGGATGACCGCAATFCAGGACTAGAGTTTAACTGCAGCGGAACATGTCTTCTCTATT 140
Db M.TH...TS.YRAA.CWN.KTCS..H..TT.H.M...R..S.RH.R....ABTTH.TEHT 279
      :|:::| :|:::| :|:::| :|:::| :|:::| :|:::| :|:::| :|:::| :|:::|
Qy 141 TCTGTGACGCGCCCCCATGCCCCCCCACCAAGAGGAGGAGGCGCCCGCATGAC 200
Db .N.MH.BTRBSMBCSMA.MSBS..HS...TB.RASKMBABH.DH...C.TBBSMB..X 219
      :|:::| :|:::| :|:::| :|:::| :|:::| :|:::| :|:::| :|:::| :|:::|
Qy 201 CCTCGTCTCTGCAGGCTGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db NS.T.DY..YY.M.S.STM.T..H.G..TKYTTA....AST..Y..C.MAYRASMR.CD. 159
      :|:::| :|:::| :|:::| :|:::| :|:::| :|:::| :|:::| :|:::| :|:::|
Qy 261 CACATGAAAAATAAAGCG 278
Db 158 A.S.NCTRD.BBAR..CB 141
      :|:::| :|:::| :|:::| :|:::| :|:::| :|:::| :|:::| :|:::| :|:::|
```

Search completed: December 6, 2005, 22:57:09
Job time : 734.068 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:34:11 ; Search time 200.691 Seconds
(without alignments)
455.919 Million cell updates/sec

Title: US-09-980-046B-7
Perfect score: 294
Sequence: 1 cggcgatgcggcggttat.....agcgagcacctagacaaaa 294

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New:*

1: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
5: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2.*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3.*
10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	15.6	633	US-10-750-185-47850	Sequence 47850, A
2	32.2	11.0	1581	US-10-750-185-29187	Sequence 29187, A
3	28.6	9.7	1138	US-10-750-185-38374	Sequence 38374, A
4	28.6	9.7	1664	US-10-750-185-26660	Sequence 26660, A
5	28	9.5	1017	US-10-750-185-36774	Sequence 36774, A
6	28	9.5	1007	US-10-750-185-53113	Sequence 53113, A
7	28	9.5	157224	US-11-112-908-51	Sequence 51, Appl
8	28	9.5	170189	US-11-112-908-50	Sequence 50, Appl
9	27.8	9.5	795	US-10-750-185-30589	Sequence 30589, A
10	27.8	9.5	1303	US-10-750-185-57470	Sequence 57470, A
11	27.6	9.4	600	US-10-750-185-3545	Sequence 3545, Ap
12	27.6	9.4	131855	US-11-112-908-29	Sequence 29, Appl
13	27.6	9.4	143389	US-11-112-908-30	Sequence 30, Appl
14	27.6	9.4	166020	US-11-112-908-28	Sequence 28, Appl
15	27.6	9.4	171247	US-11-112-908-27	Sequence 27, Appl
16	27	9.2	1126	US-10-750-185-45111	Sequence 45111, A
17	27	9.2	1233	US-10-750-185-63362	Sequence 63362, A
18	27	9.2	1316	US-10-750-185-57351	Sequence 57351, A
19	27	9.2	2624	US-10-750-185-29398	Sequence 29398, A
20	26.8	9.1	1437	US-10-750-185-32922	Sequence 32922, A
21	26.8	9.1	2024	US-10-750-185-45936	Sequence 45936, A
22	26.8	9.1	2476	US-10-986-501-86	Sequence 86, Appl
23	26.8	9.1	165883	US-11-112-908-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1

US-10-750-185-47850
; Sequence 47850, Application US/10750185
; Publication No. US20050260603A1

GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10750,185
; PRIOR FILING DATE: 2003-12-31
; CURRENT APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47850
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Bovine 19866881311131
US-10-750-185-47850

Query Match 15.6%; Score 46; DB 6; Length 633;
Best Local Similarity 57.7%; Pred. No. 8.7e-07;
Matches 82; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy	21	TCCCATGTCCTGTCAAGAGTGCACATCACTACCTAGGAGCCACCCCGCT 80
Db	203	TCGCACGACGCTCTGTCAAAGAGTGCACATCACTAGGAGCCCTCACACCA 262
Qy	81	CTGGATGACCGAATCAGGACTAGATTAACTGCAGCGGAACATGTCATTTCTCTATT 140
Db	263	ACTGCACCGAATCAGGAGCTGGTGGTGAACGAGCTCCGCTCTGTCTCTCTCTG 322
Qy	141	TCTGTGACGGCCCGCCATGCC 162
Db	323	CCCCGCCCTGCGCCGCTCCC 344

RESULT 2

US-10-750-185-29187/c
; Sequence 29187, Application US/10750185

Sequence 57135, A
Sequence 51346, A
Sequence 1313, Ap
Sequence 63168, A
Sequence 46776, A
Sequence 28000, A
Sequence 48088, A
Sequence 4443, Ap
Sequence 45911, A
Sequence 62, Appl
Sequence 90, Appl
Sequence 2883, Ap
Sequence 2839, Ap
Sequence 35161, A
Sequence 51110, A
Sequence 57640, A
Sequence 59823, A
Sequence 62973, A
Sequence 25256, A
Sequence 36, Appl
Sequence 35, Appl

```
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29187
; LENGTH: 1581
; TYPE: DNA
; ORGANISM: Bovine 19866881898685
US-10-750-185-29187

Query Match 11.0%; Score 32.2; DB 6; Length 1581;
Best Local Similarity 59.1%; Pred. No. 0.12;
Matches 55; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy 140 TTCTGTGAGCGCCCGCCATGCCCCACCCACCAAGAGCGGAGGCGCCAGTCATGA 199
Db 769 TGCTGTGAGCACACCCCGCGCCCTCCCGCCAAAGCGGCGAGTGAGGCGCTAGTTTGG 710

Qy 200 CCCTCGTCTCTGCGAGGCTGTGTACAGTTCT 232
Db 709 GCCAGGCGCGGTAGGATTTGGACAGGGTCT 677

RESULT 3
US-10-750-185-38374
; Sequence 38374, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38374
; LENGTH: 1138
; TYPE: DNA
; ORGANISM: Bovine 19866880599576
US-10-750-185-38374

Query Match 9.7%; Score 28.6; DB 6; Length 1138;
Best Local Similarity 57.1%; Pred. No. 2;
Matches 52; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 128 TCATTTCCTATTCTGTGAGCGCGCCCGCCATGCCCCACCCACCAAGAGCGGAGG 187
Db 24 TCATTTCCTTTTATCTGACTACCCCGCCCGTCTCTGTGCCCGCCCGACAGCAGG 83

Qy 188 GCCCAGTCATGACCCCTCTCTGTGAGGGTG 218
Db 84 GCTGAGTCAATTTTCGAGGATCTCTGAGTCTG 114

; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26660
; LENGTH: 1664
; TYPE: DNA
; ORGANISM: Bovine 19866881515211
US-10-750-185-26660

Query Match 9.7%; Score 28.6; DB 6; Length 1664;
Best Local Similarity 57.1%; Pred. No. 2.4;
Matches 52; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 25 ATGATGCTCTGTCAAGGAGTGAGCAATCACTTACTGAGGAGCCCGCCGCTTGG 84
Db 1095 AGGATGCTCTAGTTCTGTGGAGGGTGCTGTCATCTTCTGTGGGCACACACTCCAGCCCCAG 1154

Qy 85 ATGACCGCAATCAGGACTAGAGTTTAACTG 115
Db 1155 TTGCACCGCATGCACGANTAGACTGAACGG 1185

RESULT 5
US-10-750-185-36774/c
; Sequence 36774, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36774
; LENGTH: 917
; TYPE: DNA
; ORGANISM: Bovine 19866880785228
US-10-750-185-36774

Query Match 9.5%; Score 28; DB 6; Length 917;
Best Local Similarity 53.7%; Pred. No. 3;
Matches 58; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 59 CTGAGGAGCCCGCCCGCTCTGATGAGCCGCAATCAGGACTAGAGTTTAACTGAG 118
```

Db 145 CTGCAAAAGCCAGGCGCGTGGTGAAGAGTACACCTGGGCGGCTCTGGCCCTGGCCTG 86
Qy 119 CGGAACATGTCTATTTCTTATTTCTGTGACGCGGCCCCCATGCCCCAC 166
Db 85 GTCAGTAGACATTGGGCGCCTGCACTGTGGGCGCCCGAGGACCACTC 38

RESULT 6

US-10-750-185-53113/c
; Sequence 53113, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53113
; LENGTH: 1007
; TYPE: DNA
; ORGANISM: Bovine 19866881478112
US-10-750-185-53113

Query Match 9.5%; Score 28; DB 6; Length 1007;
Best Local Similarity 48.2%; Pred. No. 3.1;
Matches 79; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy 23 CCATGATGCTCTGTCAAGGAGTGACAACTACCTACTGAGGAGCCCAACCCCGCTCT 82
Db 814 CCACAGTGCTGGCAGCAGCAGTCTTCTTGTAGGATGCCCACTCCCTCC 755
Qy 83 GGATGGACCGAATCAGGAGTGTAACTGCGAGGGAACATGCTATTTCTCTATTTC 142
Db 754 GGGTCCCTCTATTTCTGGGAGCTGTAATTAGACAACTTCCCGCCAGTTGCTATCTG 695
Qy 143 TGTAGCGCCCGCCATGCCCCCACCACCCACCAAGAGGAGGAG 186
Db 694 TGGCAGCAGTCACTATTCCTCAGGCGCTGCGCCACTGCCACTAG 651

RESULT 7

US-11-112-908-51
; Sequence 51, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51
; LENGTH: 157224

; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-51
Query Match 9.5%; Score 28; DB 7; Length 157224;
Best Local Similarity 51.6%; Pred. No. 35;
Matches 64; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
Qy 2 GGCATGCGGCGGCGTTATTTCCTCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTG 61
Db 40234 GGTGCTGGTCTGCTGGGACTCTTGTAGTCACTTGGCTGGGAGTTGAAAGTAGCCTGAGA 40293
Qy 62 AGAGCCCCACCCCGCTCTGGATGGACCGAATCAGGACTAGAGTTTAAGTGCAGCGG 121
Db 40294 ATGTACCTCTCCAACTCTCATCTTACAGAGGCCCATGGAGGAAACTGACCTGCTAGG 40353
Qy 122 AACCA 125
Db 40354 AACCA 40357

RESULT 8

US-11-112-908-50
; Sequence 50, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 170189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-50

Query Match 9.5%; Score 28; DB 7; Length 170189;
Best Local Similarity 51.6%; Pred. No. 36;
Matches 64; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 2 GGCATGCGGCGGCGTTATTTCCTCATGATGCTCTGTCAAGGAGTGGACAATCACCTACTG 61
Db 72994 GGTGCTGGTCTGCTGGGACTCTTGTAGTCACTTGGCTGGGAGTTGAAAGTAGCCTGAGA 73053
Qy 62 AGAGCCCCACCCCGCTCTGGATGGACCGAATCAGGACTAGAGTTTAAGTGCAGCGG 121
Db 73054 ATGTACCTCTCCAACTCTCATCTTACAGAGGCCCATGGAGGAAACTGACCTGCTAGG 73113
Qy 122 AACCA 125
Db 73114 AACCA 73117

RESULT 9

US-10-750-185-30589/c
; Sequence 30589, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard

Ds 93358 GTCATTCAGCCTGGTGACAGAGTGAGATCCTGTCTATAAAAAAAAAAAGGCACTCT 93417
Qy 262 ACATGAAAAATAAGCGA 279
Db 93418 GCATCCAAAAAGGAGA 93435

RESULT 13

US-11-112-908-30
; Sequence 30, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; PRIOR FILING DATE: 2005-04-22
; PRIOR FILING DATE: 2004-04-23
; PRIOR FILING DATE: 2004-06-01
; PRIOR FILING DATE: 2004-06-01
; PRIOR FILING DATE: 2004-11-30
; PRIOR FILING DATE: 2004-11-30
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 30
; LENGTH: 143389
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-30

Query Match 9.4%; Score 27.6; DB 7; Length 143389;

Best Local Similarity 50.0%; Pred. No. 47;
Matches 69; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 142 CTGTGACGCGCCCGCCATGCCCGCCACACAGAGGCGAGGCGCCAGTCATGACC 201
Db 61071 CAGTAGCTCGCGCCCGCCATGCCCGCCACAGAGCTGAGGTGGGAGGATTGCTTAA 61130
Qy 202 CTCGTCTCTGACGGGTGTCTGTACAGCTTCTCTGAAAGCTTTAGACAGTAACAGTTTGC 261
Db 61131 GTCATTCAGCCTGGTGACAGTGAGATCCTGTCTATAAAAAAAAAAAGGCACTCT 61190
Qy 262 ACATGAAAAATAAGCGA 279
Db 61191 GCATCCAAAAAGGAGA 61208

RESULT 14

US-11-112-908-28
; Sequence 28, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR FILING DATE: 2004-04-23
; PRIOR FILING DATE: 2004-06-01
; PRIOR FILING DATE: 2004-06-01
; PRIOR FILING DATE: 2004-11-30
; PRIOR FILING DATE: 2004-11-30
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 28

; LENGTH: 166020
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-28

Query Match

Best Local Similarity 50.0%; Pred. No. 50;
Matches 69; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 142 CTGTGACGCGCCCGCCATGCCCGCCACACAGAGGCGAGGCGCCAGTCATGACC 201
Db 107552 CAGTAGCTCGCGCCCGCCATGCCCGCCACAGAGCTGAGGTGGGAGGATTGCTTAA 107611
Qy 202 CTCGTCTCTGACGGGTGTCTGTACAGCTTCTCTGAAAGCTTTAGACAGTAACAGTTTGC 261
Db 107612 GTCATTCAGCCTGGTGACAGTGAGATCCTGTCTATAAAAAAAAAAAGGCACTCT 107671
Qy 262 ACATGAAAAATAAGCGA 279
Db 107672 GCATCCAAAAAGGAGA 107689

RESULT 15

US-11-112-908-27
; Sequence 27, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR FILING DATE: 2005-04-22
; PRIOR FILING DATE: 2004-04-23
; PRIOR FILING DATE: 2004-06-01
; PRIOR FILING DATE: 2004-06-01
; PRIOR FILING DATE: 2004-11-30
; PRIOR FILING DATE: 2004-11-30
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 27
; LENGTH: 171247
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-27

Query Match 9.4%; Score 27.6; DB 7; Length 171247;

Best Local Similarity 50.0%; Pred. No. 51;
Matches 69; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 142 CTGTGACGCGCCCGCCATGCCCGCCACACAGAGGCGAGGCGCCAGTCATGACC 201
Db 157511 CAGTAGCTCGCGCCCGCCATGCCCGCCACAGAGCTGAGGTGGGAGGATTGCTTAA 157570
Qy 202 CTCGTCTCTGACGGGTGTCTGTACAGCTTCTCTGAAAGCTTTAGACAGTAACAGTTTGC 261
Db 157571 GTCATTCAGCCTGGTGACAGTGAGATCCTGTCTATAAAAAAAAAAAGGCACTCT 157630
Qy 262 ACATGAAAAATAAGCGA 279
Db 157631 GCATCCAAAAAGGAGA 157648

Search completed: December 6, 2005, 23:16:54
Job time : 202.691 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:57 ; Search time 3355.79 Seconds
(without alignments)
4099.008 Million cell updates/sec

Title: US-09-980-046B-7
Perfect score: 294
Sequence: 1 cggcgatcgcgcgcttat.....agcgagcacctagacaaaa 294

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_hcc: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_est7: *
9: gb_gss1: *
10: gb_gss2: *
11: gb_gss3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	271.4	92.3	670	BM470718	AGENCOURT
2	270.8	92.1	330	AA804321	AA804321 nw26911.s
3	270.8	92.1	346	BQ690721	AGENCOURT
4	270.8	92.1	397	AA599515	ag26e07.s
5	270.8	92.1	411	BQ018738	UI-H-DH1-
6	270.8	92.1	419	BQ652690	AGENCOURT
7	270.8	92.1	470	AA599195	ag34a07.s
8	270.8	92.1	575	BQ008059	UI-H-E10-
9	270.8	92.1	588	BUE32242	UI-H-FE1-
10	270.8	92.1	753	CA431219	UI-H-FG1-
11	270.4	92.0	428	AA156259	zo4Bb04.r
12	269.4	91.6	543	CR736990	CR736990
13	269.4	91.6	628	CK430066	CK430066 oJ44f10.y
14	269.4	91.6	628	CM428169	170005999
15	269.4	91.6	650	CB114667	CB114667 K-EST0158
16	269.4	91.6	692	AV718358	AV718358
17	269.4	91.6	723	CR990355	CR990355
18	269.4	91.6	740	BG881135	BG881135 602828965
19	269.4	91.6	809	BG169773	BG169773 602324595
20	269.4	91.6	874	CV810559	AGENCOURT
21	269.4	91.6	895	CV811638	AGENCOURT
22	269.4	91.6	997	BM458195	AGENCOURT

23	269.4	91.6	1038	3	BM543106	BM543106
24	269.4	91.6	1095	3	BM456796	AGENCOURT
25	269.4	91.6	1542	4	CR595687	full-leng
26	269.4	91.6	1761	4	HSB800862	AL110208 Homo sapi
27	269.4	91.6	2785	4	HSB800418	AL050108 Homo sapi
28	269.2	91.6	391	3	BM670864	BM670864 UI-E-DH1-
29	269.2	91.6	459	3	BM991388	BM991388 UI-H-DH0-
30	269.2	91.6	466	6	CA437724	CA437724 UI-H-DH0-
31	269.2	91.6	468	1	AI281760	AI281760 qt86g12.x
32	269	91.5	1185	1	AL563732	AL563732
33	268.8	91.4	899	5	BQ691167	AGENCOURT
34	268.4	91.3	717	7	CT002650	CT002650
35	267.8	91.1	488	3	BM832046	BM832046 K-EST0106
36	267.8	91.1	563	3	BM759791	K-EST0040
37	267.8	91.1	754	5	BQ431810	BQ431810 AGENCOURT
38	267.8	91.1	1041	5	BX405988	BX405988
39	267.6	91.0	392	1	AI147712	AI147712 qb43c08.x
40	267.6	91.0	1076	1	AL579733	AL579733
41	267.4	91.0	368	8	244036	HSC1RB041 n
42	267	90.8	858	5	BU540159	BU540159 AGENCOURT
43	266.2	90.5	600	1	AI991096	AI991096 wu38a06.x
44	266.2	90.5	968	2	BG697854	BG697854 602661067
45	266	90.5	419	1	AA400985	AA400985 zu50e04.s

ALIGNMENTS

RESULT 1
BM470718
LOCUS
DEFINITION AGENCOURT 6463211 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5533535
5', mRNA sequence.
ACCESSION BM470718
VERSION BM470718.1 GI:18519760
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 670)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12218 row: i column: 24
High quality sequence stop: 622.
Location/Qualifiers
1. .670
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5533535"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."

FEATURES
source
Query Match 92.3%; Score 271.4; DB 3; Length 670;
Best Local Similarity 99.6%; Pred. No. 2.7e-73;
Matches 272; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
ORIGIN
Query Match 92.3%; Score 271.4; DB 3; Length 670;
Best Local Similarity 99.6%; Pred. No. 2.7e-73;
Matches 272; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 21 TCCCATGATGCTCTGTCAAGGAGTGACAAATCACCTACTGAGGAGCCCCACCCGCGCT 80
Db 387 TCCCATGATGCTCTGTCAAGGAGTGACAAATCACCTACTGAGGAGCCCCACCCGCGCT 446
QY 81 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGACGCGGAACATGTCTCTATT 140
Db 447 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGACGCGGAACATGTCTCTATT 506
QY 141 TCTGTGACGGCGCCCATGCCCCACCCACCAACAGAGGAGGAGGCCCCAGTCTATGAC 200
Db 507 TCTGTGACGGCGCCCATGCCCCACCCACCAACAGAGGAGGAGGCCCCAGTCTATGAC 566
QY 201 CCTCGTCTCTGACGGGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db 567 CCTCGTCTCTGACGGGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 626
QY 261 CACATGAAAAATAAAGCGAGCAGCCTAGACAAAA 293
Db 627 CACATGAAAAATAAAGCGAGCAGCCTAAACAAAA 659

```

RESULT 2

```

AA804321/c
LOCUS      330 bp      mRNA      linear      EST 18-FEB-1998
DEFINITION      n226g11.e1 NCI_CGAP_GCB0 Homo sapiens cDNA clone IMAGE:1241636 3',
                mRNA sequence.

```

```

ACCESSION      AA804321
VERSION        AA804321.1  GI:2873608
KEYWORDS       EST.
SOURCE         Homo sapiens (human)

```

ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

```

```

REFERENCE      1 (bases 1 to 330)

```

```

AUTHORS        NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

```

```

TITLE          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

```

```

JOURNAL        Tumor Gene Index

```

```

COMMENT        Unpublished (1997)

```

```

Contact: Robert Strausberg, Ph.D.

```

```

Email: cgapbs-remail.nih.gov

```

```

Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,

```

```

Ph.D., Gerald Marti, M.D.

```

```

CDNA Library Preparation: Louis M. Staudt, M.D., Ph.D.

```

```

DNA Sequencing by: Greg Lennon, Ph.D.

```

```

CDNA Library Arrayed by: Washington University Genome Sequencing Center

```

```

Cloned through the I.M.A.G.E. Consortium/LLNL at:

```

```

www-bio.llnl.gov/bbrp/image/image.html

```

```

Insert Length: 2436 Std Error: 0.00

```

```

Seq primer: -40m13 fwd. ET from Amersham

```

```

High quality sequence stop: 228.

```

```

Location/Qualifiers

```

FEATURES

```

source

```

```

1..330

```

```

/organism="Homo sapiens"

```

```

/mol_type="mRNA"

```

```

/db_xref="taxon:9606"

```

```

/clone="IMAGE:1241636"

```

```

/tissue_type="germinal center B-cells"

```

```

/lab_host="SOLR (kanamycin resistant)"

```

```

/clone_lib="NCI_CGAP_GCB0"

```

```

/notes="Organ: tonsil; Vector: Bluescript SK-; Site 1:

```

```

ECORI; Site 2: XhoI; Cloned unidirectionally. Primer:

```

```

Oligo dt. Germinal center B-cells Library constructed by

```

```

Dr. L. Staudt (NCI). 5' adaptor sequence: 5'

```

```

GAATTCGGCAGCAG 3' 3' adaptor sequence: 5'

```

```

CTCGAGTTTCTTTTCTTTTCTTTT 3' Average insert size: 1.1 kb."

```

ORIGIN

```

Query Match      92.1%; Score 270.8; DB 1; Length 330;
Best Local Similarity 99.3%; Pred. No. 3.6e-73;
Matches 272; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 21 TCCCATGATGCTCTGTCAAGGAGTGACAAATCACCTACTGAGGAGCCCCACCCGCGCT 80
Db 277 TCCCATGATGCTCTGTCAAGGAGTGACAAATCACCTACTGAGGAGCCCCACCCGCGCT 218
QY 81 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGACGCGGAACATGTCTCTATT 140
Db 217 CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGACGCGGAACATGTCTCTATT 158
QY 141 TCTGTGACGGCGCCCATGCCCCACCCACCAACAGAGGAGGAGGCCCCAGTCTATGAC 200
Db 157 TCTGTGACGGCGCCCATGCCCCACCCACCAACAGAGGAGGAGGCCCCAGTCTATGAC 98
QY 201 CCTCGTCTCTGACGGGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db 97 CCTCGTCTCTGACGGGTCTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 38
QY 261 CACATGAAAAATAAAGCGAGCAGCCTAGACAAAA 294
Db 37 CACATGAAAAATAAAGCGAGCAGCCTAAACAAAA 4

```

RESULT 3

```

BQ690721
LOCUS      346 bp      mRNA      linear      EST 15-JUL-2002
DEFINITION      AGENCOURT_8047310 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6207332
                5', mRNA sequence.

```

```

ACCESSION      BQ690721
VERSION        BQ690721.1  GI:21816037
KEYWORDS       EST.
SOURCE         Homo sapiens (human)

```

ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

```

```

REFERENCE      1 (bases 1 to 346)

```

```

AUTHORS        NIH-MGC http://mgc.nci.nih.gov/.

```

```

TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)

```

```

JOURNAL        Unpublished (1999)

```

```

COMMENT        Contact: Robert Strausberg, Ph.D.

```

```

Email: cgapbs-remail.nih.gov

```

```

Tissue Procurement: ATCC

```

```

CDNA Library Preparation: Rubin Laboratory

```

```

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```

```

DNA Sequencing by: Agencourt Bioscience Corporation

```

```

Cloned through the I.M.A.G.E. Consortium/LLNL at:

```

```

http://image.llnl.gov

```

```

Plate: LLCM2363 row: d column: 21

```

```

High quality sequence stop: 345.

```

```

Location/Qualifiers

```

FEATURES

```

source

```

```

1..346

```

```

/organism="Homo sapiens"

```

```

/mol_type="mRNA"

```

```

/db_xref="taxon:9606"

```

```

/clone="IMAGE:6207332"

```

```

/tissue_type="ductal carcinoma, cell line"

```

```

/lab_host="DH10B (phage-resistant)"

```

```

/clone_lib="NIH MGC 110"

```

```

/notes="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;

```

```

Site 2: EcoRI; cDNA made by oligo-dT priming.

```

```

Directionally cloned into EcoRI/XhoI sites using the

```

```

following 5' adaptor: GGCAGAG(G). Library constructed by

```

```

Ling Hong in the laboratory of Gerald M. Rubin (University

```

```

of California, Berkeley) using ZAP-cDNA synthesis kit

```

```

(Stratagene) and Superscript II RT (Life Technologies).

```

```

Note: this is a NIH_MGC Library."

```

ORIGIN

```

Query Match      92.1%; Score 270.8; DB 5; Length 346;
Best Local Similarity 99.3%; Pred. No. 3.6e-73;
Matches 272; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```
QY 21 TCCCATGATGCTCTGTCAAGAGTGGACAAATACCTACTGAGGAGCCCCACCCCGGCT 80
Db 62 TCCCATGATGCTCTGTCAAGAGTGGACAAATACCTACTGAGGAGCCCCACCCCGGCT 121
QY 81 CTGGATGACCGAATCAGGACTAGAGTTTAACCTGACGGGAACATGTCATTCTCTATT 140
Db 122 CTGGATGACCGAATCAGGACTAGAGTTTAACCTGACGGGAACATGTCATTCTCTATT 181
QY 141 TCTGTGACGCGCCCCCATGCCCCACCCACCAAGAGGAGGAGGCCCAAGTCATGAC 200
Db 182 TCTGTGACGCGCCCCCATGCCCCACCCACCAAGAGGAGGAGGCCCAAGTCATGAC 241
QY 201 CCTGCTCTCTGACGGGTCTGTACAGTCTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db 242 CCTGCTCTCTGACGGGTCTGTACAGTCTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 301
QY 261 CACATGAAAAATAAAGCGAGCACCTAGACAAAA 294
Db 302 CACATGAAAAATAAAGCGAGCACCTAAAAA 335
```

```
RESULT 4
AA599515/c
LOCUS AA599515 397 bp mRNA linear EST 20-JUN-2002
DEFINITION ag26e07.s1 Human bone marrow stromal cells Homo sapiens cDNA clone
IMAGE:1090692 3', mRNA sequence.
```

```
ACCESSION AA599515
```

```
VERSION AA599515.1 GI:2433140
```

```
KEYWORDS EST.
```

```
SOURCE Homo sapiens (human)
```

```
ORGANISM
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
Hominidae; Homo.
```

```
1 (bases 1 to 397)
```

```
Jia,L., Robey,P., Young,M., Schuler,G., Powell,J., Yang,L.,
```

```
Lennon,G., Hillier,L., Allen,M., Bowles,L., Geisel,S., Kucaba,T.,
```

```
Marra,M., Martin,J., Steptoe,M., Tan,F., Theising,B., Bowers,Y.,
```

```
Wylie,T., Waterston,R., Wilson,R. and Francomano,C.
```

```
WashU-MGB/NHGRI EST Project
```

```
Unpublished (1997)
```

```
Contact: Wilson RK / Jia L
```

```
WashU-MGB/NHGRI EST Project
```

```
Washington University School of Medicine
```

```
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
```

```
Tel: 314 286 1800
```

```
Fax: 314 286 1810
```

```
Email: est@watson.wustl.edu
```

```
This clone is available royalty-free through LNL ; contact the
```

```
IMAGE Consortium (info@image.llnl.gov) for further information.
```

```
Insert Length: 722 Std Error: 0.00
```

```
Seq primer: -40m13 fwd. ET from Amersham
```

```
High quality sequence stop: 391.
```

```
Location/Qualifiers
```

```
1..397
```

```
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
```

```
/clone="IMAGE:1090692"
```

```
/sex="mixed"
```

```
/tissue_type="bone marrow stroma"
```

```
/dev_stage="mixed"
```

```
/lab_host="XLI-Blue MRF"/SOLR"
```

```
/clone_lib="Human bone marrow stromal cells"
```

```
/note="Vector: pBluescript; Site 1: EcoRI; Site 2: XhoI;
```

```
mRNA made from human bone marrow stroma, cDNA made by
```

```
oligo-dT priming. Directionally cloned. Size-selected for
```

```
average insert size >0.5 kb. Library constructed by Dr.
```

```
Marian Young and Dr. Pamela Gehron Robey (NIDCR). Library
```

```
supplied by Dr. Libin Jia (NHGRI)"
```

```
ORIGIN
```

```
Query Match 92.1%; Score 270.8; DB 1; Length 397;
```

```
Best Local Similarity 99.3%; Pred. No. 3.7e-73;
Matches 272; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 21 TCCCATGATGCTCTGTCAAGAGTGGACAAATACCTACTGAGGAGCCCCACCCCGGCT 80
Db 275 TCCCATGATGCTCTGTCAAGAGTGGACAAATACCTACTGAGGAGCCCCACCCCGGCT 216
QY 81 CTGGATGACCGAATCAGGACTAGAGTTTAACCTGACGGGAACATGTCATTCTCTATT 140
Db 215 CTGGATGACCGAATCAGGACTAGAGTTTAACCTGACGGGAACATGTCATTCTCTATT 156
QY 141 TCTGTGACGCGCCCCCATGCCCCACCCACCAAGAGGAGGAGGCCCAAGTCATGAC 200
Db 155 TCTGTGACGCGCCCCCATGCCCCACCCACCAAGAGGAGGAGGCCCAAGTCATGAC 96
QY 201 CCTGCTCTCTGACGGGTCTGTACAGTCTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db 95 CCTGCTCTCTGACGGGTCTGTACAGTCTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 36
QY 261 CACATGAAAAATAAAGCGAGCACCTAGACAAAA 294
Db 35 CACATGAAAAATAAAGCGAGCACCTAAAAA 2
```

```
RESULT 5
```

```
BQ018738/c
```

```
LOCUS BQ018738 411 bp mRNA linear EST 27-MAR-2002
```

```
DEFINITION UI-H-DH1-awu-b-18-0-UI.s1 NCI CGAP_DH1 Homo sapiens cDNA clone
```

```
IMAGE:5823665 3', mRNA sequence.
```

```
ACCESSION BQ018738
```

```
VERSION BQ018738.1 GI:19754007
```

```
KEYWORDS EST.
```

```
SOURCE Homo sapiens (human)
```

```
ORGANISM
```

```
Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
Hominidae; Homo.
```

```
1 (bases 1 to 411)
```

```
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
```

```
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
```

```
Tumor Gene Index
```

```
Unpublished (1997)
```

```
Contact: Robert Strausberg, Ph.D.
```

```
Email: cgapbs-remail.nih.gov
```

```
Tissue Procurement: Dr. Jose Mercuende
```

```
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
```

```
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
```

```
Clone Distribution: Clone distribution information can be found
```

```
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
```

```
Seq primer: M13 FORWARD
```

```
POLYA=Yes.
```

```
Location/Qualifiers
```

```
1..411
```

```
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
```

```
/clone="IMAGE:5823665"
```

```
/tissue_type="Metastatic Chondrosarcoma"
```

```
/dev_stage="Adult"
```

```
/lab_host="DH10B (Life Technologies)"
```

```
/clone_lib="NCI CGAP DH1"
```

```
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
```

```
modified polylinker; Site 1: EcoR I; Site 2: Not I;
```

```
NCI CGAP DH1 is a normalized cDNA library containing the
```

```
following tissue(s): VS-8 Cell line from Metastatic
```

```
Chondrosarcoma in Lung. The library was constructed
```

```
according to Bonaldo, Lennon and Soares, Genome Research,
```

```
6:791-806, 1996. First strand cDNA synthesis was primed
```

```
with an oligo-dT primer containing a Not I site. Double
```

```
stranded cDNA was ligated to an EcoR I adaptor, digested
```

```
with Not I, and cloned directionally into pT73-Pac
```

```
vector. The oligonucleotide used to prime the synthesis of
```

first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGATCATGTC.

TAG_TISSUE=lung
TAG_LIB=UI-H-DH1
TAG_SEQ=AGATCATGTC

ORIGIN

```

Query Match      92.1%; Score 270.8; DB 3; Length 411;
Best Local Similarity 99.3%; Pred. No. 3.7e-73;
Matches 272; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 21 TCCCATGATGCTCTGTCAAGAGTGGACATACCTACTGAGGAGCCACCCCGCCT 80
    |||
Db 283 TCCCATGATGCTCTGTCAAGAGTGGACATACCTACTGAGGAGCCACCCCGCCT 224
    |||

QY 81 CTGGATGACCGAATCAGGACTAGAGTTTAACTGCAGCGGAAACATGTCTCTATT 140
    |||
Db 223 CTGGATGACCGAATCAGGACTAGAGTTTAACTGCAGCGGAAACATGTCTCTATT 164
    |||

QY 141 TCTGTGAGCGCCGCCATGCCCCACCCACCAAGAGGAGGAGGCGCCAGTCATGAC 200
    |||
Db 163 TCTGTGAGCGCCGCCATGCCCCACCCACCAAGAGGAGGAGGCGCCAGTCATGAC 104
    |||

QY 201 CCTCGTCTCTGCAGGGTGTCTGTACAGTCTTCTGAAAGCTTTAGACAGTAAAGTTTG 260
    |||
Db 103 CCTGTCTCTGCAGGGTGTCTGTACAGTCTTCTGAAAGCTTTAGACAGTAAAGTTTG 44
    |||

QY 261 CACATGAAAAATAAAGCGAGCAGCCTAGACAAAAA 294
    |||
Db 43 CACATGAAAAATAAAGCGAGCAGCCTAAACAAAAA 10
    |||

```

RESULT 6

```

BQ652690
LOCUS      BQ652690      419 bp      mRNA      linear      EST 15-JUL-2002
DEFINITION AGENCOURT_8488531 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6296172
            5', mRNA sequence.
ACCESSION  BQ652690
VERSION     BQ652690.1 GI:21776862
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1 (bases 1 to 419)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbe-remail.nih.gov
            Tissue Procurement: CGAP (Stanford)
            CDNA Library Preparation: Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2503 row: j column: 13
            High quality sequence stop: 418.
            Location/Qualifiers
                1..419
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:6296172"
                    /tissue_type="hepatocellular carcinoma, cell line"
                    /lab_host="DH10B (phage-resistant)"
                    /clone_lib="NIH_MGC_100"
                    /note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
                    EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                    into EcoRI/XhoI sites using the following 5' adaptor:

```

FEATURES

```

source
1..470
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1091412"
/location="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_100"
/note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:

```

ORIGIN

```

Query Match      92.1%; Score 270.8; DB 5; Length 419;
Best Local Similarity 99.3%; Pred. No. 3.7e-73;
Matches 272; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 21 TCCCATGATGCTCTGTCAAGAGTGGACATACCTACTGAGGAGCCACCCCGCCT 80
    |||
Db 134 TCCCATGATGCTCTGTCAAGAGTGGACATACCTACTGAGGAGCCACCCCGCCT 193
    |||

QY 81 CTGGATGACCGAATCAGGACTAGAGTTTAACTGCAGCGGAAACATGTCTCTATT 140
    |||
Db 194 CTGGATGACCGAATCAGGACTAGAGTTTAACTGCAGCGGAAACATGTCTCTATT 253
    |||

QY 141 TCTGTGAGCGCCGCCATGCCCCACCCACCAAGAGGAGGAGGCGCCAGTCATGAC 200
    |||
Db 254 TCTGTGAGCGCCGCCATGCCCCACCCACCAAGAGGAGGAGGCGCCAGTCATGAC 313
    |||

QY 201 CCTCGTCTCTGCAGGGTGTCTGTACAGTCTTCTGAAAGCTTTAGACAGTAAAGTTTG 260
    |||
Db 314 CCTGTCTCTGCAGGGTGTCTGTACAGTCTTCTGAAAGCTTTAGACAGTAAAGTTTG 373
    |||

QY 261 CACATGAAAAATAAAGCGAGCAGCCTAGACAAAAA 294
    |||
Db 374 CACATGAAAAATAAAGCGAGCAGCCTAAACAAAAA 407
    |||

```

RESULT 7

```

AA599195/c
LOCUS      AA599195      470 bp      mRNA      linear      EST 20-JUN-2002
DEFINITION AG34407.81 Human bone marrow stromal cells Homo sapiens cDNA clone
            IMAGE:1091412 3' similar to WP:C04F6.4 CE03924 BETA TRANSDUCIN ;,
            mRNA sequence.
ACCESSION  AA599195
VERSION     AA599195.1 GI:2432820
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1 (bases 1 to 470)
            Jia,L., Robey,P., Young,M., Schuler,G., Powell,J., Yang,L.,
            Lennon,G., Hillier,L., Allen,M., Bowles,L., Geisler,S., Kucaba,T.,
            Marra,M., Martin,J., Steptoe,M., Tan,F., Theising,B., Bowers,Y.,
            Wyllie,T., Waterston,R., Wilson,R. and Francimano,C.
            WashU-MGB/NHGRI EST Project
            Unpublished (1997)
            Contact: Wilson RK / Jia L
            WashU-MGB/NHGRI EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Insert Length: 2461 Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 321.
            Location/Qualifiers
                1..470
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:1091412"

```

FEATURES

```

source
1..470
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1091412"

```

GGCAGCAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

POLYA=Yes.		Location/Qualifiers		Tumor Gene Index	
source		1. .598		Unpublished (1997)	
		/organism="Homo sapiens"		Contact: Robert Strausberg, Ph.D.	
		/mol_type="mRNA"		Email: cgapbs-r@mail.nih.gov	
		/db_xref="taxon:9606"		Tissue Procurement: James Martin	
		/clone="UI-H-FGI-bab-e-07-0-UI"		cDNA Library preparation: Dr. M. Bento Soares, University of Iowa	
		/tissue_type="Cell lines"		cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa	
		/dev_stage="Adult"		Cloning Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu	
		/lab_host="DH10B (Life Technologies)"		Seq primer: M13 FORWARD	
		/clone_lib="NCI CGAP_FGI"		POLYA=Yes.	
		/notes="Organ: Chondrosarcoma; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FGI is a normalized cDNA library derived from a pool of mRNA obtained from 3 cell lines from grade II chondrosarcoma tissues. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGTACGGAC. The cell lines were provided by Dr James Martin from the University of Iowa."			
		TAG_TISSUE=Human grade 2 chondrosarcoma cell line pool			
		TAG_LIB=UI-H-FGI			
		TAG_SEQ=CGCTACGGAC"			
ORIGIN					
Query Match		92.1%; Score 270.8; DB 5; Length 588;			
Best Local Similarity		99.3%; Pred. No. 4e-73; 2; Indels 0; Gaps 0;			
Matches		272; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	21	TCCCATGATGCCTCTGTCAAGGAGTGACAAATCACCTACTGAGGAGCCACCCCGCCT	80		
Db	283	TCCCATGATGCCTCTGTCAAGGAGTGACAAATCACCTACTGAGGAGCCACCCCGCCT	224		
Qy	81	CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGAGCGGAACATGCTATTCTCTATT	140		
Db	223	CTGGATGGACCGAATCAGGAGCTAGAGTTTAACTGAGCGGAACATGCTATTCTCTATT	164		
Qy	141	TCTGTGACGGCCCGCATGCCCCACCCACCAAGAGGAGGAGGCGCCAGTCATGAC	200		
Db	163	TCTGTGACGGCCCGCATGCCCCACCCACCAAGAGGAGGAGGCGCCAGTCATGAC	104		
Qy	201	CCTCGTCTCTGACGGGTGTCTGTACAGTCTTCTGAAAGCTTTAGACAGTAACAGTTTG	260		
Db	103	CCTCGTCTCTGACGGGTGTCTGTACAGTCTTCTGAAAGCTTTAGACAGTAACAGTTTG	44		
Qy	261	CACATGAAAAATAAAGCGAGCACCTAGACAAAAA	294		
Db	43	CACATGAAAAATAAAGCGAGCACCTAGACAAAAA	10		
RESULT 10					
CA431219/c		753 bp mRNA linear EST 07-NOV-2002			
LOCUS		UI-H-FGI-bgi-1-18-0-UI.s1 NCI CGAP_FGI Homo sapiens cDNA clone			
DEFINITION		UI-H-FGI-bgi-1-18-0-UI 3', mRNA sequence.			
ACCESSION		CA431219			
VERSION		CA431219.1			
KEYWORDS		EST.			
SOURCE		Homo sapiens (human)			
ORGANISM		Homo sapiens			
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE		1 (bases 1 to 753)			
AUTHORS		NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.			
TITLE		National Cancer Institute, Cancer Genome Anatomy Project (CGAP).			
JOURNAL					
COMMENT					


```
VERSION AAI5259.1 GI:1727877
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chissole, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W.,
Hawkins, M., Kulkarni, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.
and Marra, M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
PUBMED 8889549
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LINL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham.
FEATURES
source
Location/Qualifiers
1..428
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:4621481"
/db_xref="taxon:9606"
/clone="IMAGE:590095"
/dev_stage="umbilical vein, 1 passage"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene endothelial cell 937223"
/notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; Cloned unidirectionally. Primer: Oligo dt.
Umbilical vein endothelial cells, passaged once. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
sequence: 5' GAATTCGACGAG 3' ~3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTTTTTT 3'"
ORIGIN
Query Match 92.0%; Score 270.4; DB 1; Length 428;
Best Local Similarity 99.6%; Pred. No. 5e-73;
Matches 271; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 21 TCCCATGATGCTCTGTCAAGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 80
Db 272 TCCCATGATGCTCTGTCAAGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 213
QY 81 CTGGATGACCAATCAGGAGCTAGAGTTTAACTGCAGCGGAACATGTCATTTCTCTATT 140
Db 212 CTGGATGACCAATCAGGAGCTAGAGTTTAACTGCAGCGGAACATGTCATTTCTCTATT 153
QY 141 TCTGTGACGCGCCCATGCGCCCAACCCCAACCAAGAGCGCAGAGGCGCCAGTCATGAC 200
Db 152 TCTGTGACGCGCCCATGCGCCCAACCCCAACCAAGAGCGCAGAGGCGCCAGTCATGAC 93
QY 201 CCTCGTCTCTGACGGGTGCTGTACACGTTCTTCTGAAAGCTTTTAGACAGTAACAGTTTG 260
Db 92 CCTCGTCTCTGACGGGTGCTGTACACGTTCTTCTGAAAGCTTTTAGACAGTAACAGTTTG 33
QY 261 CACATGAAAAATAAGCGAGCACCTAGACAAA 292
Db 32 CACATGAAAAATAAGCGAGCACCTAGACAAA 1
RESULT 12
CR736990
LOCUS CR736990 543 bp mRNA linear EST 10-DEC-2004
DEFINITION CR736990 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:108671 5', mRNA sequence.
ACCESSION CR736990
VERSION CR736990
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 543)
Ebert, L., Heil, O., Hennig, S., Korn, B., Neubert, P., Partsch, E.,
Peters, M., Radelof, U. and Schneider, D.
I.M.A.G.E. cDNA Clone Collection
Unpublished (2004)
Contact: Inge Arlt
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Email: www.rzpd.de
RZPD; IMAGE:998E2480.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Contact: Inge Arlt
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACAGGAACACAGCTAATGAC.
FEATURES
source
Location/Qualifiers
1..543
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:998E2480 ; IMAGE:108671"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/notes="Organ: Liver and Spleen; Vector: p773D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5', AACTGGAAGAAATTAATAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN
Query Match 91.6%; Score 269.4; DB 7; Length 543;
Best Local Similarity 99.6%; Pred. No. 1.1e-72;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 21 TCCCATGATGCTCTGTCAAGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 80
Db 218 TCCCATGATGCTCTGTCAAGAGTGGACAATCACCTACTGAGGAGCCCAACCCCGCCT 277
QY 81 CTGGATGACCAATCAGGAGCTAGAGTTTAACTGCAGCGGAACATGTCATTTCTCTATT 140
Db 278 CTGGATGACCAATCAGGAGCTAGAGTTTAACTGCAGCGGAACATGTCATTTCTCTATT 337
QY 141 TCTGTGACGCGCCCATGCGCCCAACCCCAACCAAGAGCGCAGAGGCGCCAGTCATGAC 200
Db 338 TCTGTGACGCGCCCATGCGCCCAACCCCAACCAAGAGCGCAGAGGCGCCAGTCATGAC 397
QY 201 CCTCGTCTCTGACGGGTGCTGTACACGTTCTTCTGAAAGCTTTTAGACAGTAACAGTTTG 260
Db 398 CCTCGTCTCTGACGGGTGCTGTACACGTTCTTCTGAAAGCTTTTAGACAGTAACAGTTTG 457
QY 261 CACATGAAAAATAAGCGAGCACCTAGACAAA 291
```

[illegible]

Job time : 3358.89 secs

```
RESULT 15
CB114667
LOCUS      CB114667                650 bp    mRNA    linear    EST 28-JAN-2003
DEFINITION K-EST0158371 L10choCK0 Homo sapiens cDNA clone L10choCK0-1-C03 5',
            mRNA sequence.
ACCESSION  CB114667
VERSION    CB114667.1
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE  1 (bases 1 to 650)
AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE     21C Frontier Korean EST Project 2001
JOURNAL   Unpublished (2002)
COMMENT   Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 1 row; C column: 03
            High quality sequence stop: 650.
FEATURES             Location/Qualifiers
     source           1..650
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="L10choCK0-1-C03"
                     /sex="M"
                     /cell_line="Cho-CK"
                     /lab_host="Top10P"
                     /clone_lib="L10choCK0"
                     /note="Organ: Liver; Vector: p7T3-Pac; Site_1: EcoRI;
                     Site_2: NotI; The library was contributed by the Soares
                     laboratory and it was constructed as described by Bonaldo,
                     M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
                     6(9): 791-806. RNA was prepared from harvested cell
                     culture."
ORIGIN
Query Match      91.6%; Score 269.4; DB 6; Length 650;
Best Local Similarity 99.6%; Pred. No. 1.1e-72;
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY  21  TCCCATGATGCTCTGTCAAGAGTGGACATCACCTACTGAGGAGCCCCACCCCGCCT 80
Db   |||||||
QY  236 TCCCATGATGCTCTGTCAAGAGTGGACATCACCTACTGAGGAGCCCCACCCCGCCT 295
Db   |||||||
QY  81  CTGGATGACCGCAATCAGGACTAGAGTTTAATCGCAGCGGAACATGTCTTCTCTATT 140
Db   |||||||
QY  296 CTGGATGACCGCAATCAGGACTAGAGTTTAATCGCAGCGGAACATGTCTTCTCTATT 355
Db   |||||||
QY  141 TCTGTGACGCGCCCCCATGCCCCACCCACCACCAAGAGGCGAGGCGCCAGTCATGAC 200
Db   |||||||
QY  356 TCTGTGACGCGCCCCCATGCCCCACCCACCACCAAGAGGCGAGGCGCCAGTCATGAC 415
Db   |||||||
QY  201 CCTCGTCTCTGAGGGGTGTGTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 260
Db   |||||||
QY  416 CCTCGTCTCTGAGGGGTGTGTGTACACGTTCTTCTGAAAGCTTTAGACAGTAACAGTTTG 475
Db   |||||||
QY  261 CACATGAAAAATAAAGCGGACCTAGACAA 291
Db   |||||||
QY  476 CACATGAAAAATAAAGCGGACCTAGACAA 506
Db   |||||||
```

Search completed: December 7, 2005, 09:55:57

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 21:43:03 ; Search time 910.919 Seconds
(without alignments)
6427.437 Million cell updates/sec

Title: US-09-980-046B-8

Perfect score: 103

Sequence: 1 cggacagacggcagcagtc.....tgagcccttcctcacaaaa 103

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb ba.*

2: gb in.*

3: gb env.*

4: gb om.*

5: gb ov.*

6: gb pat.*

7: gb ph.*

8: gb pr.*

9: gb ro.*

10: gb sts.*

11: gb sy.*

12: gb un.*

13: gb vl.*

14: gb btg.*

15: gb pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101.4	98.4	427	6	BD223503 94 human
2	101.4	98.4	427	6	AR243137 Sequence
3	101.4	98.4	427	6	AR404318 Sequence
4	101.4	98.4	427	6	AR576159 Sequence
5	101.4	98.4	1874	6	CQ769657 Sequence
6	101.4	98.4	2031	6	BD016437 Gene deri
7	101.4	98.4	2031	8	AF135600 Homo sapi
8	101.4	98.4	2034	8	BC037974 Homo sapi
9	101.4	98.4	2057	8	AX288046 Sequence
10	100.4	97.5	1277	6	BD223416 94 human
11	100.4	97.5	1277	6	AR243050 Sequence
12	100.4	97.5	1277	6	AR404231 Sequence
13	100.4	97.5	1277	6	AR576072 Sequence
14	98.4	95.5	136328	8	AC004126 Human Chr
15	98.4	95.5	177073	14	AF000934 Homo sapi
c 16	98.4	95.5	195616	14	AC090344 Homo sapi
c 17	98.4	95.5	197156	14	AC090345 Homo sapi
c 18	98.4	95.5	197856	8	AP000777 Homo sapi

c	19	98.4	95.5	201460	8	AP003721	AP003721 Homo sapi
	20	96.4	93.6	2113	8	AB020598	AB020598 Homo sapi
	21	93	90.3	4563	6	CQ850121	CQ850121 Sequence
	22	93	90.3	4563	8	AK127216	AK127216 Homo sapi
	23	90.4	87.8	855	10	BV445087	BV445087 HSPASBP1
	24	83.2	80.8	374	10	G29419	G29419 human STS S
	25	74.4	72.2	280	6	BD139841	BD139841 Compounds
	26	74.4	72.2	280	6	AR202945	AR202945 Sequence
	27	74.4	72.2	280	6	AR208046	AR208046 Sequence
	28	74.4	72.2	280	6	AX429919	AX429919 Sequence
	29	74.4	72.2	2786	6	CQ413062	CQ413062 Sequence
	30	73.4	71.3	454	6	BD154667	BD154667 Primer fo
	31	73.4	71.3	454	6	AX874605	AX874605 Sequence
	32	73.4	71.3	2235	6	BD160044	BD160044 Primer fo
	33	73.4	71.3	2235	6	AX882976	AX882976 Sequence
	34	73.4	71.3	2235	8	AK023577	AK023577 Homo sapi
	35	72.8	70.7	2755	8	AK000546	AK000546 Homo sapi
	36	71	68.9	1447	8	AK130141	AK130141 Homo sapi
	37	68	66.0	1902	6	AX302027	AX302027 Sequence
	38	65	63.1	245	10	G30209	G30209 human STS S
	39	63.8	61.9	317	6	AX336583	AX336583 Sequence
	40	63.8	61.9	317	6	AX336875	AX336875 Sequence
	41	60.2	58.4	189589	14	AC146641	AC146641 Ofotemur
	42	52.2	50.7	258748	14	AC164281	AC164281 Bos tauru
	43	44.8	43.5	603	10	BV266907	BV266907 SZ35P636R
	44	43.4	42.1	1979	9	AB026665	AB026665 Rattus no
	45	42.4	41.2	2042	9	BC051940	BC051940 Mus muscu

ALIGNMENTS

RESULT 1	BD223503	427 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	94 human secretory proteins.				
DEFINITION	BD223503				
ACCESSION	BD223503.1	GI:33033273			
VERSION	JP 2002518010-A/106.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 427)				
AUTHORS	Ruben, S.M., Ni, J., Rosen, C.A., Wei, Y.F., Young, P.E., Florence, K.A., Soppet, D.R., Brewer, L.A., Endress, G.A., Carter, K.C., Mucenski, M., Ebner, R., Lafleur, D.W., Olsen, H.S., Shi, Y., Moore, P.A. and Komatsoulis, G.				
TITLE	94 human secretory proteins				
JOURNAL	Patent: JP 2002518010-A 106 25-JUN-2002;				
COMMENT	HUMAN GENOME SCIENCES INC				
	OS Homo sapiens (human)				
	PN JP 2002518010-A/106				
	PD 25-JUN-2002				
	PF 15-JUN-1999 JP 2000554850				
	PR 16-JUN-1998 US 60/089509,16-JUN-1998 US 60/089510 PR				
	16-JUN-1998 US 60/089508,16-JUN-1998 US 60/089507 PR				
	22-JUN-1998 US 60/090112,22-JUN-1998 US 60/090113 PI STEVEN M RUBEN, JIAN NI, CRAIG A ROSEN, YING FEI WEI, PAUL E YOUNG,				
	PI KIMBERLY A FLORENCE, DANIEL R Soppet, LAURIE A BREWER, GREGORY A PI ENDRESS,				
	PI KENNETH C CARTER, MICHAEL MUCENSKI, REINHARD EBNER, DAVID W PI LAFLEUR,				
	PI HENRIK S OLSEN, YANGGU SHI, PAUL A MOORE, GEORGE KOMATSOUKIS PC C12N15/09, A61K38/00, A61K39/395, A61K39/395, A61K48/00, A61P43/00, PC C07K7/00,				
	PC C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC , C12Q3/68, C12N15/00,				
	PC A61K37/02, C12N5/00				
	CC 94 human secretory proteins				
	Key Location/Qualifiers				

FT source 1. .427
FT Location/Qualifiers
FEATURES
source 1. .427
/organism="Homo sapiens (human)"
/db_xref="taxon:9606"
ORIGIN
Query Match 98.4%; Score 101.4; DB 6; Length 427;
Best Local Similarity 99.0%; Pred. No. 4e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGACAGCGGACGAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 60
DB 281 CGGACAGCGGACGAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 340
QY 61 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 103
DB 341 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 383
RESULT 2
AR243137
LOCUS AR243137 427 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 107 from patent US 6475753.
ACCESSION AR243137
VERSION AR243137.1 GI:27289830
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 427)
AUTHORS Ruben,S.M., Ni,J., Rosen,C.A., Wei,Y.-F., Young,P., Florence,K., Soppet,D.R., Brewer,L.A., Endress,G.A., Carter,K.C., Mucenski,M., Ebner,R., Lafleur,D.W., Olsen,H., Shi,Y., Moore,P.A. and Komatsoulis,G.
TITLE 94 Human Secreted Proteins
JOURNAL Patent: US 6475753-A 107 05-NOV-2002;
Human Genome Sciences, Inc.; Rockville, MD
FEATURES
source 1. .427
Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 98.4%; Score 101.4; DB 6; Length 427;
Best Local Similarity 99.0%; Pred. No. 4e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGACAGCGGACGAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 60
DB 281 CGGACAGCGGACGAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 340
QY 61 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 103
DB 341 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 383
RESULT 3
AR404318
LOCUS AR404318 427 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 107 from patent US 6627741.
ACCESSION AR404318
VERSION AR404318.1 GI:40152416
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 427)
AUTHORS Ruben,S.M., Ni,J., Rosen,C.A., Wei,Y.-F., Young,P., Florence,K., Soppet,D.R., Brewer,L.A., Endress,G.A., Carter,K.C., Mucenski,M., Ebner,R., Lafleur,D.W., Olsen,H., Shi,Y., Moore,P.A. and Komatsoulis,G.
TITLE 94 Human Secreted Proteins
JOURNAL Patent: US 6627741-A 107 10-SEP-2003;
Human Genome Sciences, Inc.; Rockville, MD
FEATURES
source 1. .427
Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 98.4%; Score 101.4; DB 6; Length 427;
Best Local Similarity 99.0%; Pred. No. 4e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGACAGCGGACGAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 60
DB 281 CGGACAGCGGACGAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 340
QY 61 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 103
DB 341 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 383
RESULT 4
AR576159
LOCUS AR576159 427 bp DNA linear PAT 14-DEC-2004
DEFINITION Sequence 107 from patent US 6774216.
ACCESSION AR576159
VERSION AR576159.1 GI:56577705
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 427)
AUTHORS Ruben,S.M., Rosen,C.A. and Lafleur,D.W.
TITLE Antibodies to secreted protein HCEJQ69
JOURNAL Patent: US 6774216-A 107 10-AUG-2004;
Human Genome Sciences, Inc.; Rockville, MD
FEATURES
source 1. .427
Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 98.4%; Score 101.4; DB 6; Length 427;
Best Local Similarity 99.0%; Pred. No. 4e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGACAGCGGACGAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 60
DB 281 CGGACAGCGGACGAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 340
QY 61 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 103
DB 341 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 383
RESULT 5
CQ769657
LOCUS CQ769657 1874 bp DNA linear PAT 04-MAR-2004
DEFINITION Sequence 383 from Patent WO2003058021.
ACCESSION CQ769657
VERSION CQ769657.1 GI:45114200
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Koening-Hoffman,K., Kazinski,M., Schaefer,R. and Kesper,B.
TITLE Novel apoptosis-inducing dna sequences
JOURNAL Patent: WO 2003058021-A 383 17-JUL-2003;

Komatsoulis,G.
Antibodies to secreted protein HCEJQ69
Patent: US 6627741-A 107 30-SEP-2003;
Human Genome Sciences, Inc.; Rockville, MD
FEATURES
source 1. .427
Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 98.4%; Score 101.4; DB 6; Length 427;
Best Local Similarity 99.0%; Pred. No. 4e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGACAGCGGACGAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 60
DB 281 CGGACAGCGGACGAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 340
QY 61 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 103
DB 341 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 383
RESULT 4
AR576159
LOCUS AR576159 427 bp DNA linear PAT 14-DEC-2004
DEFINITION Sequence 107 from patent US 6774216.
ACCESSION AR576159
VERSION AR576159.1 GI:56577705
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 427)
AUTHORS Ruben,S.M., Rosen,C.A. and Lafleur,D.W.
TITLE Antibodies to secreted protein HCEJQ69
JOURNAL Patent: US 6774216-A 107 10-AUG-2004;
Human Genome Sciences, Inc.; Rockville, MD
FEATURES
source 1. .427
Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 98.4%; Score 101.4; DB 6; Length 427;
Best Local Similarity 99.0%; Pred. No. 4e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGACAGCGGACGAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 60
DB 281 CGGACAGCGGACGAGTCCAGCTCTGGTTCTCTCGGTTTATTCTGTAGATGAA 340
QY 61 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 103
DB 341 ATGGTTCCCATTAATAAGGGGCATGAGCCCTTCTCACAATAA 383
RESULT 5
CQ769657
LOCUS CQ769657 1874 bp DNA linear PAT 04-MAR-2004
DEFINITION Sequence 383 from Patent WO2003058021.
ACCESSION CQ769657
VERSION CQ769657.1 GI:45114200
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Koening-Hoffman,K., Kazinski,M., Schaefer,R. and Kesper,B.
TITLE Novel apoptosis-inducing dna sequences
JOURNAL Patent: WO 2003058021-A 383 17-JUL-2003;

FEATURES
source
Xantos Biomedicine AG (DE)
Location/Qualifiers
1. .1874
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 98.4%; Score 101.4; DB 6; Length 1874;
Best Local Similarity 99.0%; Pred. No. 3.5e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTATTCGTTAGATGAA 60
Dy 1707 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTATTCGTTAGATGAA 1766
Qy 61 ATGGTTCCCATAAATAGGGGCATGAGCCCTTCCTCACAAAA 103
Dy 1767 ATGGTTCCCATAAATAGGGGCATGAGCCCTTCCTCACAAAA 1809

RESULT 6
BD016437
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
OS
PN
PD
PF
PI
PC
PC
PC
CC
FH
FT
FEATURES
source
Xantos Biomedicine AG (DE)
Location/Qualifiers
1. .1874
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 98.4%; Score 101.4; DB 6; Length 2031;
Best Local Similarity 99.0%; Pred. No. 3.5e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTATTCGTTAGATGAA 60
Dy 1909 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTATTCGTTAGATGAA 1968
Qy 61 ATGGTTCCCATAAATAGGGGCATGAGCCCTTCCTCACAAAA 103
Dy 1969 ATGGTTCCCATAAATAGGGGCATGAGCCCTTCCTCACAAAA 2011

RESULT 7
AF135600
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
AF135600
Homo sapiens osteoclast transporter hoCTP mRNA
cde.
AF135600
AF135600.1 GI:33337870
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 2031)
Yamane,S., Toyosaki-Maeda,T., Tsuruta,Y., Suzuki,R. and Ochi,T.
Cloning of a novel gene of peptide transporter from human
osteoclast
Unpublished
2 (bases 1 to 2031)
Yamane,S., Toyosaki-Maeda,T., Tsuruta,Y., Suzuki,R. and Ochi,T.
Direct Submission
Submitted (19-MAR-1999) Shionogi Institute for Medical Science,
Shionogi & Co., Ltd., 2-5-1 Mishima, Settu, Osaka 566-0022, Japan
Shionogi
Location/Qualifiers
1. .2031
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q12.2"
/cell_type="osteoclast"
/tissue_type="RA synovial fluid"
1. .2031
/gene="OCTP"
128. .1873
/gene="OCTP"
/note="genomic sequence found within Homo sapiens
chromosome 11q12.2 PAC clone pJ60696 deposited in Genbank
Accession Number AC004126"
/codon_start=1
/product="osteoclast transporter hoCTP"
/protein_id="AAQ13565.1"
/db_xref="GI:33337871"
/translation="MPAPRAREQPRVPERQPLPRGARGRRWRRAAGAAVLLVEML
ERAAFFGVGTANLVLYNSTNFTNQEQATRAALVFLGASVLLAPVGGMLADVILGRYR
AVALSLLLYLAASGLLPATAPPDGRSSFCGEMPASPLGPAKPCSPSPSPYCAPV
LYAGLLLLGAAASVRSNLATSRGADQVMDLGRDATRFFNFYWSINLGASVLLVVA
FIQNIIFLLGYSIPVCGVGLAFIFLFPVFTKPPMSQVSMKLALQNCPPQL
WQHSARDQCARVLADERSPPQASPOEDIANFQVILPVMVTLVPMVTFPMQ
STVYLQGLHLHPNIPFANPANTSVALLRAQSSYTIPEAWLLANVVVLLVLPKDR
LIDPLLRCKLLPSALQKMGFMFGFTSVIVAGVLEMERLHYTHNETHVQQIGEV
YNAAPLSIWWQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQGAIMGIFPCLSGVGS
LGSSVALLSLPGGMLHCPKDPFGNINNCMDLYFFLLAGIQAVTALLFWIAGRYERA
SQGPASHSRFSRDRG"
2008
/gene="OCTP"
polyA_site

ORIGIN
Query Match 98.4%; Score 101.4; DB 8; Length 2031;
Best Local Similarity 99.0%; Pred. No. 3.5e-23;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTATTCGTTAGATGAA 60
Dy 1909 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTATTCGTTAGATGAA 1968
Qy 61 ATGGTTCCCATAAATAGGGGCATGAGCCCTTCCTCACAAAA 103
Dy 1969 ATGGTTCCCATAAATAGGGGCATGAGCCCTTCCTCACAAAA 2011
RESULT 8
BC037974

BC037974 2034 bp mRNA linear PRI 30-JUN-2004
 HOMO sapiens solute carrier family 15, member 3, mRNA (cDNA clone MGC:46085 IMAGE:5213632), complete cds.

BC037974
 VERSION BC037974.1 GI:23273712
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 2034)
 Strausberg R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Maitland, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 12477932

2 (bases 1 to 2034)
 Strausberg, R.
 Direct Submission
 Submitted (20-SEP-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgaps@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@hgrl.nih.gov
 Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masello, C., Mabeiri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Teurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 78 Row: 9 Column: 11
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7706116.
 Location/Qualifiers
 1..2034
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:46085 IMAGE:5213632"
 /tissue_type="Blood, adult leukocytes"

LOCUS BC037974 2034 bp mRNA linear PRI 30-JUN-2004
 DEFINITION HOMO sapiens solute carrier family 15, member 3, mRNA (cDNA clone MGC:46085 IMAGE:5213632), complete cds.
 ACCESSION BC037974
 VERSION BC037974.1 GI:23273712
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 2034)
 Strausberg R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Maitland, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 12477932

2 (bases 1 to 2034)
 Strausberg, R.
 Direct Submission
 Submitted (20-SEP-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgaps@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@hgrl.nih.gov
 Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masello, C., Mabeiri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Teurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 78 Row: 9 Column: 11
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7706116.
 Location/Qualifiers
 1..2034
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:46085 IMAGE:5213632"
 /tissue_type="Blood, adult leukocytes"

clone_lib="NIH MGC_118"
 /lab_host="DH10B"
 /notes="Vector: pCMV-SPORT6"
 1..2034
 /genes="SLC15A3"
 /notes="synonyms: PTR3, PHT2, hPTR3"
 /db_xref="GeneID:51296"
 95..1840
 /genes="SLC15A3"
 /codon_start=1
 /product="solute carrier family 15, member 3"
 /protein_id="AAH37974.1"
 /db_xref="GI:23273713"
 /db_xref="GeneID:51296"
 /translation="MPARAREQPRVPGERQPLPRGARGPRRRRAAGAAVLLVEMLE
 ERAAFGVTVANLVLYLNTFNWGTQATRAALVFLGASVLLAPVGVGLADVILGRYR
 AVALLSLLYLAASGLLPATAPFDGRSFCGEMPASPLGACPSAGPRSPSPYCAPV
 LVAGLLLLGLAASSVRSNLTSFGADQMDLGRDARRFFNFWYMSNLGALVSLVVA
 FIQONTSPILGVSIPIVCGVLAFFIIEATPVTITKPPMGSOVSSMLKALONCCPOL
 WQRHSARDQCARVLADERSPOGASPOEDIANFOVLKILPVMTLVPMVYFOMQ
 STYVLOGHLHLPNFFPANPANSVALRAQGSYTIPEAWLLLANVVVLLVPLKDR
 LIDPLLRCKLLPSALOKWALGNFFGFTSVIVAGVLEMERLHYIHNETVSQIQIGVL
 YNAAPLSIWMQIPQYLLIGISEIFASIPGLEPAYSEAPRSMQGAIMGIFPCLSGVGL
 LGSSVALLSLPGGMLRCPKDFGNNCRMDLYFFLLAGIQVATALLFVMIAGRYERA
 SQGPASHSRFSRDRG"

ORIGIN

Query Match 98.4%; Score 101.4; DB 8; Length 2034;
 Best Local Similarity 99.0%; Pred. No. 3.5e-23;
 Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGCGGCAGTCCAGCTCTGGTTCTCTCGGTTATCTCTAGAAATGAA 60
 |||||
 DB 1876 CGGACAGCGGCAGTCCAGCTCTGGTTCTCTCGGTTATCTCTAGAAATGAA 1935
 |||||

QY 61 ATGGTTCCCAATAAAGGGGCATGAGCCCTCTCACAAAA 103
 |||||
 DB 1936 ATGGTTCCCAATAAAGGGGCATGAGCCCTCTCACAAAA 1978
 |||||

RESULT 9
 LOCUS AX288046 2057 bp DNA linear PAT 21-NOV-2001
 DEFINITION Sequence 1 from Patent WO0179497.
 ACCESSION AX288046
 VERSION AX288046.1 GI:17049786
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.

REFERENCE 1
 AUTHORS Curtiss, R.A.
 TITLE 33556, a human transporter and uses thereof
 JOURNAL Patent: WO 0179497-A 1 25-OCT-2001;
 Millennium Pharmaceuticals (US)
 FEATURES Location/Qualifiers
 source 1..2057
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="cDNA of human WTP"

ORIGIN

Query Match 98.4%; Score 101.4; DB 6; Length 2057;
 Best Local Similarity 99.0%; Pred. No. 3.5e-23;
 Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGCGGCAGTCCAGCTCTGGTTCTCTCGGTTATCTCTAGAAATGAA 60
 |||||
 DB 1914 CGGACAGCGGCAGTCCAGCTCTGGTTCTCTCGGTTATCTCTAGAAATGAA 1973
 |||||

Qy 61 ATGTTCCATAAATAAGGGGCATGAGCCCTTCTCACAATAA 103
|||||
Db 1974 ATGTTCCATAAATAAGGGGCATGAGCCCTTCTCAGAAA 2016
|||||

RESULT 10
BD223416 LOCUS 1277 bp DNA linear PAT 17-JUL-2003
DEFINITION 94 human secretory proteins.
ACCESSION BD223416
VERSION BD223416.1 GI:33033186
KEYWORDS JP 2002518010-A/19.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 1277)
Ruben,S.M., Ni,J., Rosen,C.A., Wei,Y.F., Young,P.E., Florence,K.A.,
Soppet,D.R., Brewer,L.A., Endress,G.A., Carter,K.C., Mucenski,M.,
Ebner,R., Lafleur,D.W., Olsen,H.S., Shi,Y., Moore,P.A. and
Komatsoulis,G.
94 human secretory proteins
Patent: JP 2002518010-A 19 25-JUN-2002;
HUMAN GENOME SCIENCES INC
OS -Homo sapiens (human)
PN JP 2002518010-A/19
PD 25-JUN-2002
PF 15-JUN-1999 JP 2000554850
PR 16-JUN-1998 US 60/089509,16-JUN-1998 US 60/089510 PR
16-JUN-1998 US 60/089508,16-JUN-1998 US 60/089507 PR
22-JUN-1998 US 60/090112,22-JUN-1998 US 60/090113 PI STEVEN
M RUBEN, JIAN NI, CRAIG A ROSEN, YING
FEI WEI, PAUL E YOUNG
PI KIMBERLY A FLORENCE, DANIEL R SOPPET, LAURIE A BREWER, GREGORY A
PI KENNETH C CARTER, MICHAEL MUCENSKI, REINHARD EBNER, DAVID W PI
LAFLEUR,
PI HENRIK S OLSEN, YANGGU SHI, PAUL A MOORE, GEORGE KOMATSOU LIS PC
C12N15/09, A61K38/00, A61K39/395, A61K39/395, A61K48/00, A61P43/00, PC
C07K7/00,
PC C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
, C12Q1/68, C12N15/00,
PC A61K37/02, C12N5/00
CC n equals a,t,g, or c
CC n equals a,t,g, or c
FH Key Location/Qualifiers
FT source 1..1277
FT /organism='Homo sapiens (human)'.
FEATURES
source Location/Qualifiers
1..1277
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN
Query Match 97.5%; Score 100.4; DB 6; Length 1277;
Best Local Similarity 98.1%; Pred. No. 7.8e-23;
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATTCGTAGAAATGAA 60
|||||
Db 1130 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATTCGTAGAAATGAA 1189
|||||

Qy 61 ATGTTCCATAAATAAGGGGCATGAGCCCTTCTCACAATAA 103
|||||
Db 1190 ATGTTCCATAAATAAGGGGCATGAGCCCTTCTCACAATAA 1232
|||||

RESULT 11
AR243050 LOCUS 1277 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 20 from patent US 6475753.

ACCESSION AR243050
VERSION AR243050.1 GI:27289743
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1277)
AUTHORS Ruben,S.M., Ni,J., Rosen,C.A., Wei,Y.-F., Young,P., Florence,K.,
Soppet,D.R., Brewer,L.A., Endress,G.A., Carter,K.C., Mucenski,M.,
Ebner,R., Lafleur,D.W., Olsen,H., Shi,Y., Moore,P.A. and
Komatsoulis,G.
94 Human Secreted Proteins
Patent: US 6475753-A 20 05-NOV-2002;
Human Genome Sciences, Inc.; Rockville, MD
FEATURES
source Location/Qualifiers
1..1277
/organism='unknown'
/mol_type='genomic DNA'
ORIGIN
Query Match 97.5%; Score 100.4; DB 6; Length 1277;
Best Local Similarity 98.1%; Pred. No. 7.8e-23;
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATTCGTAGAAATGAA 60
|||||
Db 1130 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATTCGTAGAAATGAA 1189
|||||

Qy 61 ATGTTCCATAAATAAGGGGCATGAGCCCTTCTCACAATAA 103
|||||
Db 1190 ATGTTCCATAAATAAGGGGCATGAGCCCTTCTCACAATAA 1232
|||||

RESULT 12
AR404231 LOCUS 1277 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 20 from patent US 6627741.
ACCESSION AR404231
VERSION AR404231.1 GI:40152329
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1277)
AUTHORS Ruben,S.M., Ni,J., Rosen,C.A., Wei,Y.-F., Young,P., Florence,K.,
Soppet,D.R., Brewer,L.A., Endress,G.A., Carter,K.C., Mucenski,M.,
Ebner,R., Lafleur,D.W., Olsen,H., Shi,Y., Moore,P.A. and
Komatsoulis,G.
94 Human Secreted Proteins
Patent: US 6627741-A 20 30-SEP-2003;
Human Genome Sciences, Inc.; Rockville, MD
FEATURES
source Location/Qualifiers
1..1277
/organism='unknown'
/mol_type='genomic DNA'
ORIGIN
Query Match 97.5%; Score 100.4; DB 6; Length 1277;
Best Local Similarity 98.1%; Pred. No. 7.8e-23;
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATTCGTAGAAATGAA 60
|||||
Db 1130 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATTCGTAGAAATGAA 1189
|||||

Qy 61 ATGTTCCATAAATAAGGGGCATGAGCCCTTCTCACAATAA 103
|||||
Db 1190 ATGTTCCATAAATAAGGGGCATGAGCCCTTCTCACAATAA 1232
|||||

RESULT 13
AR576072 LOCUS 1277 bp DNA linear PAT 14-DEC-2004

DEFINITION Sequence 20 from patent US 6774216.
 ACCESSION AR576072.1 GI:56577618
 VERSION AR576072.1 GI:56577618
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1277)
 AUTHORS Ruben,S.M., Rosen,C.A. and LaFleur,D.W.
 TITLE Antibodies to secreted protein HCEJQ69
 JOURNAL Patent: US 6774216-A 20 10-AUG-2004;
 Human Genome Sciences, Inc.; Rockville, MD
 FEATURES
 source Location/Qualifiers
 1..1277
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Query Match 97.5%; Score 100.4; DB 6; Length 1277;
 Best Local Similarity 98.1%; Pred.No. 7.8e-23;
 Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 CGCAGACGCGCAGCAGCTCCAGCTCTGGTTTCCTTCGCGTTTATTCCTGTAGATGAA 60
 Db 1130 CGCAGACGCGCAGCAGCTCCAGCTCTGGTTTCCTTCGCGTTTATTCCTGTAGATGAA 1189
 Qy 61 ATGTTTCCCAATAATAGGGCGATGAGCCCTTCCTCAAAAA 103
 Db 1190 ATGTTTCCCAATAATAGGGCGATGAGCCCTTCCTCAAAAA 1232
 RESULT 14
 AC004126 136328 bp DNA linear PRI 20-JUN-1998
 LOCUS Human Chromosome 11q12.2 PAC clone PDJ60696, complete sequence.
 DEFINITION
 ACCESSION AC004126
 VERSION AC004126.1 GI:3242744
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 REFERENCE 1 (bases 1 to 136328)
 AUTHORS Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M.,
 Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J.,
 English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
 Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N.,
 Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,
 Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,
 Ward,T. and Wilson,R.
 TITLE HTGS Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 136328)
 AUTHORS Evans,G.A., Athanasiou,M., Basit,M., Bradbury,P., Brignac,S.,
 Bumeister,R., Davis,C., English,C., Franklin,T.L., Garner,H.R.,
 Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J.,
 Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P.,
 Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,
 Ward,T. and Wilson,R.
 TITLE Direct Submission
 JOURNAL Submitted (06-FEB-1998) Genome Science & Technology Center,
 University of Texas Southwestern Medical Center, 5323 Harry Hines
 Blvd, Dallas, TX 75235-8591, USA
 REFERENCE 3 (bases 1 to 136328)
 AUTHORS Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M.,
 Buettner,J., Bumeister,R., Card,P., deSailboat,F., Dunn,J.,
 English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
 Grant,O., Hahner,L., Harris,J., Lewis,E., Loo,H., Loo,K.N.,
 Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,
 Schageman,J., Schilling,P., Schultz,R., Stinson,S., Syed,M. and Ward,T.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUN-1998) Genome Science & Technology Center,
 University of Texas Southwestern Medical Center, 5323 Harry Hines

COMMENT
 Blvd, Dallas, TX 75235-8591, USA
 On Jun 20, 1998 this sequence version replaced gi:2842778.
 IMPORTANT: This submission contains the entire insert of clone
 PDJ60696. PDJ60696 comes from a PAC library constructed at the
 Roswell Park Cancer Institute by the Pieter de Jong group. This
 clone has been finished according to strict quality criteria and
 attempts have been made to resolve all base calling problems such
 as compressions and repetitive elements. The expected Phred/Phrap
 calculated errors/10kb is 0.44. In addition, attempts have been
 made to assure over 99% of consensus base calls consist of either
 double-stranded coverage or 2 types of labeling chemistry on one
 strand.
 CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome
 11p12.2 Best's disease region mapped between STS D11S461 and BSR
 AHNK. This region spans over 1.5 Mbp.
 MARKER CONFIRMATION: Sequence confirmed STSs; D11S461, SHGC-36412,
 WI-12296
 MAPPED CLONE OVERLAP: HTGS phase 1 PAC clone pDJ1081b4.

FEATURES
 source Location/Qualifiers
 1..136328
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 3..111
 /rpt_family="Alu"
 repeat_region 4399..4574
 /rpt_family="Alu"
 repeat_region 5447..5700
 /rpt_family="Alu"
 repeat_region 6418..6708
 /rpt_family="Alu"
 repeat_region 8866..9147
 /rpt_family="Alu"
 repeat_region 10302..10604
 /rpt_family="Alu"
 repeat_region 10827..11119
 /rpt_family="Alu"
 repeat_region complement(12208..12498)
 /rpt_family="Alu"
 repeat_region complement(13578..13749)
 /rpt_family="MER39"
 repeat_region complement(13647..13737)
 /rpt_family="MER21"
 repeat_region complement(13805..13947)
 /rpt_family="MER39"
 repeat_region complement(13822..13908)
 /rpt_family="MER21"
 repeat_region complement(14283..14759)
 /rpt_family="MLT1"
 repeat_region 14798..14988
 /rpt_family="MER20"
 repeat_region complement(15469..15753)
 /rpt_family="Alu"
 repeat_region complement(16650..16934)
 /rpt_family="Alu"
 repeat_region complement(18976..19291)
 /rpt_family="Alu"
 repeat_region 19340..19651
 /rpt_family="Alu"
 repeat_region 20030..20309
 /rpt_family="Alu"
 repeat_region 20780..21041
 /rpt_family="Alu"
 repeat_region complement(23675..23976)
 /rpt_family="Alu"
 repeat_region complement(24041..24140)
 /rpt_family="MER"
 repeat_region complement(25164..25433)
 /rpt_family="Alu"
 repeat_region 26572..26622
 /rpt_family="MER5"
 repeat_region 27020..27310
 /rpt_family="Alu"

```

repeat_region 28214..28492
/rpt_family="Alu"
complement(30334..30612)
repeat_region 106147..106441
/rpt_family="MIR"
complement(97690..97823)
repeat_region 108079..108291
/rpt_family="Alu"
complement(97690..97823)
repeat_region 108358..108646
/rpt_family="L1"
complement(97690..97823)
repeat_region 108658..109058
/rpt_family="Alu"
complement(97690..97823)
repeat_region 109810..109942
/rpt_family="L1"
complement(97690..97823)

Query Match 95.5% Score 98.4; DB 8; Length 136328;
Best Local Similarity 99.0%; Pred.No. 2.5e-22;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGACGGCAGCAGCTCCAGCTCTGGTTTCCTTCGCGTTTATTCGTGTAAGTAA 60
Db 37299 CGGACAGACGGCAGCAGCTCCAGCTCTGGTTTCCTTCGCGTTTATTCGTGTAAGTAA 37358

Qy 61 ATGTTTCCCATTAATAAGGGCATGAGCCCTTCCTCAAA 100
Db 37359 ATGTTTCCCATTAATAAGGGCATGAGCCCTTCCTCAAA 37398

RESULT 15
AP000934 177073 bp DNA linear HTG 30-MAY-2000
LOCUS Homo sapiens chromosome 11 clone RP11-804B24 map 11q13, WORKING
DEFINITION DRAFT SEQUENCE, 43 unordered pieces.
ACCESSION AP000934
VERSION AP000934.2 GI:8119075
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 177073)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 177,073 genomic DNA of 11q13
Published Only in DataBase (1999)
TITLE Homo sapiens 177,073 genomic DNA of 11q13
JOURNAL Published Only in DataBase (1999)
REFERENCE 2 (bases 1 to 177073)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (17-DEC-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagaminara, Kanagawa 228-8555,
Japan (E-mail:hattori@gsc.riken.go.jp,
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)
On May 31, 2000 this sequence version replaced gi:6997748.
COMMENT ----- Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: Humdraft11
Center clone name: RP11-804B24
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 151872 bases at least Q40
Consensus quality: 164006 bases at least Q30
Consensus quality: 169903 bases at least Q20
Insert size: 172873; sum-of-contigs

```

Quality coverage: 4.18x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 43 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```
1 18996 contig of 18996 bp in length
19097 30931 contig of 11835 bp in length
31032 41792 contig of 10167 bp in length
41893 52059 contig of 9375 bp in length
52160 61534 contig of 5028 bp in length
61635 66662 contig of 8125 bp in length
66763 74887 contig of 6299 bp in length
74988 81286 contig of 6577 bp in length
81387 88063 contig of 4795 bp in length
88064 92858 contig of 5078 bp in length
92959 98036 contig of 4083 bp in length
98137 102220 contig of 4039 bp in length
102320 107130 contig of 4811 bp in length
107231 111269 contig of 4039 bp in length
111270 111369 contig of 4039 bp in length
111370 114943 contig of 3574 bp in length
114944 115044 contig of 3574 bp in length
115044 120076 contig of 5033 bp in length
120077 120176 contig of 5033 bp in length
120177 123973 contig of 3797 bp in length
123974 124073 contig of 3797 bp in length
124074 125880 contig of 3321 bp in length
125881 129227 contig of 3347 bp in length
129228 129327 contig of 3321 bp in length
129328 132748 contig of 3010 bp in length
132749 135758 contig of 3010 bp in length
135759 138433 contig of 2575 bp in length
138434 138533 contig of 2575 bp in length
138534 140537 contig of 2004 bp in length
140538 140637 contig of 2004 bp in length
140638 142449 contig of 1812 bp in length
142450 142549 contig of 1812 bp in length
142550 145260 contig of 2711 bp in length
145261 145360 contig of 2711 bp in length
145361 148934 contig of 3574 bp in length
148935 149034 contig of 3574 bp in length
149035 151360 contig of 2326 bp in length
151361 151460 contig of 2326 bp in length
151461 153788 contig of 2328 bp in length
153789 156021 contig of 2133 bp in length
156022 156121 contig of 2133 bp in length
156122 157966 contig of 1845 bp in length
157967 158066 contig of 1845 bp in length
158067 159715 contig of 1649 bp in length
159716 159815 contig of 1649 bp in length
159816 161440 contig of 1625 bp in length
161441 161540 contig of 1625 bp in length
161541 163065 contig of 1525 bp in length
163066 163165 contig of 1525 bp in length
163166 164694 contig of 1529 bp in length
164695 164794 contig of 1529 bp in length
164795 166644 contig of 1850 bp in length
166645 166744 contig of 1850 bp in length
166745 167983 contig of 1239 bp in length
167984 168083 contig of 1239 bp in length
168084 169363 contig of 1280 bp in length
169364 169463 contig of 1280 bp in length
169464 169581 contig of 118 bp in length
169582 169681 contig of 118 bp in length
169682 170795 contig of 1114 bp in length
170796 170895 contig of 1114 bp in length
170896 172522 contig of 1627 bp in length
172523 172622 contig of 1627 bp in length
172623 174311 contig of 1689 bp in length
174312 174411 contig of 1689 bp in length
174412 175674 contig of 1263 bp in length
175675 175774 contig of 1263 bp in length
175775 177073 contig of 1299 bp in length
177073 177073 contig of 1299 bp in length.
```

Sequence updated (26-May-2000).
NOTE: This is a 'working draft' sequence. It currently consists of 43 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 18996: contig of 18996 bp in length
18997 19096: gap of 100 bp
19097 30931: contig of 11835 bp in length
30932 31032: gap of 100 bp
31032 41792: contig of 10167 bp in length
41793 41893: gap of 100 bp
41893 52059: contig of 9375 bp in length
52060 52160: gap of 100 bp
52160 61534: contig of 5028 bp in length
61535 61635: gap of 100 bp
61635 66662: contig of 5028 bp in length
66663 66762: gap of 100 bp
```

```
FEATURES
  source      Location/Qualifiers
1..177073
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="11"
  /map="11q13"
  /clone="RP11-804B24"
misc_feature  1..18996
  /note="assembly_fragment"
misc_feature  19097..30931

Query Match      95.5%; Score 98.4; DB 14; Length 177073;
Best Local Similarity 99.0%; Pred. No. 2.4e-22;
Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 CGGACAGACGGCAGCAGTCCCGAGCTCTGGTTTCCTTCTCGGTTTATTCTGTAGAAATGAA 60
      |||
Db      30399 CGGACAGACGGCAGCAGTCCCGAGCTCTGGTTTCCTTCTCGGTTTATTCTGTAGAAATGAA 30458
      |||

Qy      61 ATGGTTCCCATATAAGGGGCATGAGCCCTTCTCTCAAA 100
      |||
Db      30459 ATGGTTCCCATATAAGGGGCATGAGCCCTTCTCTCAGA 30498
      |||
```

Search completed: December 7, 2005, 04:22:11
Job time : 913.919 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 161.924 Seconds
(without alignment)
4239.411 Million cell updates/sec

Title: US-09-980-046B-8
Perfect score: 103
Sequence: 1 cggacagcagcgacgagtc.....tgagcccttcctcacaaaaa 103

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002as.*
- 7: geneseqn2002bs.*
- 8: geneseqn2003as.*
- 9: geneseqn2003bs.*
- 10: geneseqn2003cs.*
- 11: geneseqn2003ds.*
- 12: geneseqn2004as.*
- 13: geneseqn2004bs.*
- 14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	100.0	103	4 AAC89739	AAC89739 Human gas
2	102.6	99.6	935	6 AAD34122	AAD34122 Human sec
3	102.2	99.2	1147	6 AAD34143	AAD34143 Human sec
4	102.2	99.2	1149	6 ABL89938	ABL89938 Human pol
5	101.4	98.4	427	3 ADA297115	ADA297115 Human sec
6	101.4	98.4	427	8 ADA40310	ADA40310 Human sec
7	101.4	98.4	427	9 ACH66744	ACH66744 Novel hum
8	101.4	98.4	427	9 ADB911346	ADB911346 Human sec
9	101.4	98.4	439	6 AAD34144	AAD34144 Human sec
10	101.4	98.4	1148	4 AAK56586	AAK56586 Human imm
11	101.4	98.4	1874	10 ADI62940	ADI62940 Human apo
12	101.4	98.4	2031	5 AAI68202	AAI68202 Human ost
13	101.4	98.4	2057	13 ADQ88225	ADQ88225 Human 335
14	101.4	98.4	2060	6 AAI70679	AAI70679 Human mem
15	101.4	98.4	2124	4 AAK52347	AAK52347 Human pol
16	100.4	97.5	1277	3 AAZ97028	AAZ97028 Human sec
17	100.4	97.5	1277	8 ADA39759	ADA39759 Human sec
18	100.4	97.5	1277	9 ACH66657	ACH66657 Novel hum
19	100.4	97.5	1277	9 ADB91127	ADB91127 Human sec

20	98.4	95.5	10143	4	AAK68371	AAK68371 Human imm
21	98.4	95.5	26555	4	AAK68605	AAK68605 Human imm
22	98.4	95.5	26555	4	AAK68372	AAK68372 Human imm
23	98.4	95.5	26555	4	AAI62833	AAI62833 Human gen
24	98.4	95.5	136328	6	ABZ35015	ABZ35015 Human gen
25	96.4	93.6	2113	10	ADD18985	ADD18985 Human dis
26	96.4	93.6	2113	13	ADR25796	ADR25796 Breast ca
27	96.4	93.6	2113	13	ADR14202	ADR14202 Human NF-
28	96.4	93.6	2113	13	ADR44036	ADR44036 Human kid
29	96.4	93.6	2113	13	ADP25102	ADP25102 PRO polyp
30	95.8	93.0	207	2	AAT23231	AAT23231 Human gen
31	93	90.3	4563	13	ADR07084	ADR07084 Full leng
32	90.4	87.8	1104	4	ABA08944	ABA08944 Human pep
33	90.4	87.8	1104	4	AAK53331	AAK53331 Human pol
34	90.4	87.8	1522	10	ADE09920	ADE09920 Novel DNA
35	74.8	72.6	6397	5	ABA09671	ABA09671 Human bon
36	74.4	72.2	157	2	AAT22763	AAT22763 Human gen
37	74.4	72.2	280	2	AAK84204	AAK84204 DNA encod
38	74.4	72.2	280	3	AAK79433	AAK79433 cDNA sequ
39	74.4	72.2	280	6	ABK28977	ABK28977 Human bre
40	74.4	72.2	2735	12	ADL72094	ADL72094 Human pro
41	74.4	72.2	2735	14	ADY97963	ADY97963 DNA encod
42	74.4	72.2	2786	5	ADL61921	ADL61921 Human ova
43	73.4	71.3	454	4	AAH12675	AAH12675 Human CDN
44	73.4	71.3	476	6	ABT10272	ABT10272 Human bre
45	73.4	71.3	478	14	ACL56078	ACL56078 Human col

ALIGNMENTS

RESULT 1
AAC89739
ID AAC89739 standard; cDNA; 103 BP.
XX
AC AAC89739;
XX
DT 12-MAR-2001 (first entry)
XX
DE Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 8.
XX
KW Human; cytostatic; immunomodulator; immunostimulant; vulnery;
KW anti-inflammatory; neuroprotective; antibacterial; gene therapy;
KW gastrointestinal inflammation; immune system disorder; genetic disorder;
KW cancer; autoimmune disorder; infection; wound healing; ss.
XX
OS Homo sapiens.
XX
FN WO200073324-A2.
XX
PD 07-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US015191.
XX
PR 01-JUN-1999; 99US-0137058P.
XX
PA (DIGI-) DIGITAL GENE TECHNOLOGIES INC.
XX
PI Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Haseel KW;
XX
DR WPI; 2001-061508/07.
XX
PT New polynucleotides and polypeptides, useful in gene therapy and in
PT diagnosing a pathological condition, e.g. for modulating gene expression
PT in gastrointestinal inflammation, or for treating cancers or genetic
XX
PS Claim 1; Page 85; 108pp; English.
XX
CC The present sequence is one of a number of isolated human polynucleotides
CC which are useful in gene therapy, and for diagnosing a pathological
CC condition or a susceptibility to it. In particular, the polynucleotides
CC are useful for modulating gene expression in gastrointestinal

CC inflammation. The polynucleotides are useful for chromosome
 CC identification, controlling gene expression through triple helix
 CC formation or antisense DNA or RNA, or identifying individuals from minute
 CC biological samples using DNA-based identification techniques. The
 CC polynucleotides can also be used as an alternative to restriction
 CC fragment length polymorphism (RFLP), by determining the actual base-by-
 CC base DNA sequences of selected portions of an individual's genome. The
 CC polynucleotides may also be used as molecular weight markers on Southern
 CC gels, as diagnostic probes for the presence of a specific mRNA, as a
 CC probe to substract-out known sequences in the process of discovering novel
 CC polynucleotides, or as an antigen to elicit an immune response. The
 CC polypeptides are useful in diagnostic procedures to detect a disorder.
 CC The polynucleotides and polypeptides are useful for preventing, treating
 CC or ameliorating immune system disorders, genetic disorders, cancers, some
 CC autoimmune disorders, or infections. The polynucleotides and polypeptides
 CC are also useful for differentiating, proliferating or attracting cells,
 CC leading to the regeneration of tissues, especially in wounds or burns.
 CC The polypeptides and polynucleotides may also be used as a food additive
 CC or preservative
 XX
 SQ Sequence 103 BP; 27 A; 26 C; 23 G; 27 T; 0 U; 0 Other;

Query Match 100.0%; Score 103; DB 4; Length 103;
 Best Local Similarity 100.0%; Pred. No. 4.7e-28;
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CGGACAGCGGACGAGTCCGAGCTCTGGTTCTCTCGGTTATTCGTTAGATGAA 60
 Db 1 CGGACAGCGGACGAGTCCGAGCTCTGGTTCTCTCGGTTATTCGTTAGATGAA 60
 Qy 61 ATGGTTCCTCCATTAATAGGGGCGATGAGCCCTTCTCACAAAAA 103
 Db 61 ATGGTTCCTCCATTAATAGGGGCGATGAGCCCTTCTCACAAAAA 103

RESULT 2
 AAD341122
 ID AAD341122 standard; cDNA; 935 BP.
 AC AAD341122;
 XX 16-JUL-2002 (first entry)
 XX Human secreted protein-encoding gene 17 cDNA clone HNHOD23, SEQ ID NO:27.
 XX Human; secreted protein; proliferative disorder; cancer; tumour; AIDS;
 KW foetal abnormality; developmental; haematopoietic disorder; kidney; skin;
 KW immune system; cardiovascular; angiogenic; gastrointestinal; endocrine;
 KW pregnancy-related disorder; cognitive; neurological; autoimmune disease;
 KW inflammation; allergy; rheumatoid arthritis; Alzheimer's disease; asthma;
 KW Parkinson's disease; schizophrenia; psoriasis; sepsis; atherosclerosis;
 KW diabetes; infection; wound healing; vulnery; chemotaxis; food additive;
 KW gene therapy; gene; ss.

OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT CDS 466..741
 FT /*tag= a
 FT /product= "Human secreted protein"
 FT sig_peptide 466..567
 FT /*tag= b
 FT mat_peptide 568..738
 FT /*tag= c
 FT /product= "Mature secreted protein"

XX WO200222654-A1.
 XX
 XX 21-MAR-2002.
 XX
 XX 17-JAN-2001; 2001WO-US001385.
 XX
 XX 12-SEP-2000; 2000US-0231969P.

XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR;
 PI Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;
 PI Fiscella M, Ni J;
 XX WPI; 2002-315684/35.
 DR P-PSDB; AAE21594.

Seventeen nucleic acid molecules encoding human secreted proteins, useful for treating and preventing cancer, immune disorders (e.g. Addison's disease, and allergies), and cardiovascular disorders (e.g. myocardial ischemias).

Claim 1; Page 423; 483pp; English.

XX AAD34106-AAD34144 represent cDNAs corresponding to 17 human secreted protein genes, and AAE21578-AAE21616 represent the proteins they encode. The genes AAE21617-AAE21627 represent human secreted protein fragments. The genes and their corresponding secreted proteins are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Pathological conditions can be diagnosed by determining the amount of the new protein in a sample or by determining the presence of mutations in the new genes. Specific uses are described for each of the 17 genes, based on the tissues in which they are most highly expressed, and include developing products for the diagnosis or treatment of cancer, proliferative disorders, tumours, foetal and developmental abnormalities, haematopoietic disorders, diseases of the immune system, AIDS, autoimmune diseases (rheumatoid arthritis), inflammation, allergies, neurological disorders (Alzheimer's disease, Parkinson's disease), cognitive disorder, schizophrenia, asthma, skin disorders (psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney disorders, gastrointestinal disorders, pregnancy-related disorders, endocrine disorders, and infections. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues, to identify their cognate ligands or binding partners, and in chemotaxis, and can be used as a food additive or preservative to modify storage properties. The present sequence represents a human secreted protein-encoding cDNA of the invention

SQ Sequence 935 BP; 237 A; 261 C; 241 G; 186 T; 0 U; 10 Other;

Query Match 99.6%; Score 102.6; DB 6; Length 935;
 Best Local Similarity 99.0%; Pred. No. 1.6e-27;
 Matches 102; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACAGCGGACGAGTCCGAGCTCTGGTTCTCTCGGTTATTCGTTAGATGAA 60

Db 777 CGGACAGCGGACGAGTCCGAGCTCTGGTTCTCTCGGTTATTCGTTAGATGAA 836

Qy 61 ATGGTTCCTCCATTAATAGGGGCGATGAGCCCTTCTCACAAAAA 103

Db 837 ATGGTTCCTCCATTAATAGGGGCGATGAGCCCTTCTCACAAAAA 879

RESULT 3
 AAD34143
 ID AAD34143 standard; cDNA; 1147 BP.

XX AAD34143;

XX 16-JUL-2002 (first entry)

XX Human secreted protein-encoding gene 17 cDNA clone HNHOD23, SEQ ID NO:48.

XX Human; secreted protein; proliferative disorder; cancer; tumour; AIDS;
 KW foetal abnormality; developmental; haematopoietic disorder; kidney; skin;
 KW immune system; cardiovascular; angiogenic; gastrointestinal; endocrine;
 KW pregnancy-related disorder; cognitive; neurological; autoimmune disease;
 KW inflammation; allergy; rheumatoid arthritis; Alzheimer's disease; asthma;

KW Parkinson's disease; schizophrenia; psoriasis; sepsis; atherosclerosis;
KW diabetes; infection; wound healing; vulnery; chemotaxis; food additive;
XX gene therapy; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
FH CDS 678..953
FT /*tag= a
FT /product= "Human secreted protein"
FT sig_peptide 678..779
FT /*tag= b
FT mat_peptide 780..950
FT /*tag= c
FT /product= "Mature secreted protein"

XX WO200222654-A1.

XX 21-MAR-2002.

XX 17-JAN-2001; 2001WO-US001385.

XX 12-SEP-2000; 2000US-0231969P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Komatoulis GA, Baker KP, Birse CE, Soppet DR;

PI Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;

PI Fiscella M, Ni J;

XX WPI; 2002-315684/35.

DR P-PSDB; AAE21615.

XX Seventeen nucleic acid molecules encoding human secreted proteins, useful
FT for treating and preventing cancer, immune disorders (e.g. Addison's
FT disease, and allergies), and cardiovascular disorders (e.g. myocardial
FT ischaemia).

XX Claim 1; Page 436; 483pp; English.

XX AAD34106-AAD34144 represent cDNAs corresponding to 17 human secreted
CC protein genes, and AAE21578-AAE21616 represent the proteins they encode.
CC AAE21617-AAE21627 represent human secreted protein fragments. The genes
CC and their corresponding secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 17 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of cancer,
CC proliferative disorders, tumours, foetal and developmental abnormalities,
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC diseases (rheumatoid arthritis), inflammation, allergies, neurological
CC disorders (Alzheimer's disease, Parkinson's disease), cognitive disorder,
CC schizophrenia, asthma, skin disorders (psoriasis), sepsis, diabetes,
CC atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney
CC disorders, gastrointestinal disorders, pregnancy-related disorders,
CC endocrine disorders, and infections. The proteins can also be used to aid
CC wound healing and epithelial cell proliferation, to prevent skin aging
CC due to sunburn, to maintain organs before transplantation, for supporting
CC cell culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties. The
CC present sequence represents a human secreted protein-encoding cDNA of the
CC invention

XX Sequence 1147 BP; 249 A; 354 C; 286 G; 251 T; 0 U; 7 Other;

Query Match 99.2%; Score 102.2; DB 6; Length 1147;
Best Local Similarity 98.1%; Pred. No. 2.5e-27;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTCTTCTCGGTTTATTCTGTAGATGAA 60

Db 989 CGGACAGACGGCAGCAGTCCAGCTCTGGTTCTTCTCGGTTTATTCTGTAGATGAA 1048

QY 61 ATGGTTCCCAATAAATAAGGGGCATGAGCCCTTCTCACAAAA 103

Db 1049 ATGGTTCCCAATAAATAAGGGGCATGAGCCCTTCTCAMRAAAA 1091

RESULT 4

ABL89938

ID ABL89938 standard; cDNA; 1149 BP.

XX ABL89938;

XX 24-MAY-2002 (first entry)

XX Human polynucleotide SEQ ID NO 500.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; gene; ss.

XX Homo sapiens.

XX WO200190304-A2.

XX 29-NOV-2001.

XX 18-MAY-2001; 2001WO-US016450.

XX 19-MAY-2000; 2000US-0205515P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-122018/16.

DR P-PSDB; ABB89529.

XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.

XX Claim 4; SEQ ID NO 500; 2081pp + Sequence Listing; English.

XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1149 BP; 250 A; 348 C; 292 G; 252 T; 0 U; 7 Other;

Query Match 99.2%; Score 102.2; DB 6; Length 1149;
Best Local Similarity 98.1%; Pred. No. 2.5e-27;
Matches 101; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTCTTCTCGGTTTATTCTGTAGATGAA 60

Db 991 CGGACAGCGCAGCAGCTCCAGCTCTGGTTTCTCTCTGTTATTCTCTTAGAATGAA 1050
|||
Qy 61 ATGTTTCCCATTAATAAGGGGCATGAGCCCTTCTCAGAAAA 103
|||
Db 1051 ATGTTTCCCATTAATAAGGGGCATGAGCCCTTCTCAGAAAA 1093
|||
RESULT 5
AAZ97115 ID AAZ97115 standard; cDNA; 427 BP.
XX AC AAZ97115;
XX 19-APR-2000 (first entry)
XX Human secreted protein gene 10 cDNA clone HDPWU34, SEQ ID NO:107.
XX Human; secreted protein; cancer; tumour; developmental abnormality;
KW foetal deficiency; blood disorder; immune system disorder; inflammation;
KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
KW schizophrasia; arthritis; aschma; psoriasis; sepsis; skin disorder;
KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
therapy; ds.
XX Homo sapiens.
XX WO9966041-A1.
XX 23-DEC-1999.
XX 15-JUN-1999; 99WO-US013418.
XX 16-JUN-1998; 98US-0089507P.
PR 16-JUN-1998; 98US-0089508P.
PR 16-JUN-1998; 98US-0089509P.
PR 16-JUN-1998; 98US-0089510P.
PR 22-JUN-1998; 98US-0090112P.
PR 22-JUN-1998; 98US-0090113P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Ni J, Rosen CA, Wei Y, Young PE, Florence KA;
PI Soppet DR, Brewer LA, Endress GA, Carter KC, Mucenski M, Ebner R;
PI Lafleur DW, Olsen HS, Shi Y, Moore PA, Komatsoulis G;
XX WPI; 2000-106100/09.
DR P-PSDB; AAY86311.
XX New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders.
XX Claim 1; Page 374; 586pp; English.
XX AAZ97019 to AAZ97137 represent 94 isolated human secreted protein genes.
CC AAY86215 to AAY86333 are the secreted proteins encoded by the 94 human
CC genes. This sequence represents a fragment of one of the human secreted
CC proteins. The genes and their corresponding secreted polypeptides are
CC useful for preventing, treating or ameliorating medical conditions, e.g.,
CC by protein or gene therapy. Also pathological conditions can be diagnosed
CC by determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new genes. Specific uses are
CC described for each of the 94 genes, based on which tissues they are most
CC highly expressed in, and include developing products for the diagnosis or
CC treatment of cancer, tumours, developmental abnormalities and foetal
CC deficiencies, blood disorders, diseases of the immune system, autoimmune
CC diseases, inflammation, allergies, Alzheimer's and cognitive disorders,
CC schizophrasia, arthritis, aschma, psoriasis, sepsis, skin disorders,
CC atherosclerosis, diabetes, cardiovascular disorders, kidney disorders,
CC digestive/endocrine disorders, infections and AIDS. The polypeptides are
CC also useful for identifying their binding partners. The sequences shown

CC in AAY86334 to AAY86585 represent fragments of the secreted proteins
XX Sequence 427 BP; 115 A; 120 C; 102 G; 90 T; 0 U; 0 Other;
SQ Query Match 98.4%; Score 101.4; DB 3; Length 427;
Best Local Similarity 99.0%; Pred. No. 3.3e-27;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCTCTCTGTTATTCTCTTAGAATGAA 60
|||
Db 281 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCTCTCTGTTATTCTCTTAGAATGAA 340
|||
Qy 61 ATGTTTCCCATTAATAAGGGGCATGAGCCCTTCTCAGAAAA 103
|||
Db 341 ATGTTTCCCATTAATAAGGGGCATGAGCCCTTCTCAGAAAA 383
|||
RESULT 6
ADA40310 ID ADA40310 standard; cDNA; 427 BP.
XX ADA40310;
AC ADA40310;
XX 20-NOV-2003 (first entry)
DT Human secreted protein encoding cDNA.
XX Human; secreted protein; cancer; hyperproliferative disorder;
KW rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;
KW anaemia; allergic reaction; asthma; cardiovascular disorder;
KW wound healing; cytostatic; immunosuppressive; neutropenic; neuroprotective;
KW antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory;
KW vulnary; cardiant; gene therapy; ss.
XX Homo sapiens.
XX OS
XX WO2002102993-A2.
XX 27-DEC-2002.
XX 19-MAR-2002; 2002WO-US008123.
XX 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
PI WPI; 2003-175238/17.
XX New human secreted proteins and nucleic acid molecules, useful for
PT preparing a diagnostic or pharmaceutical composition for diagnosing,
PT preventing or treating cancer or other hyperproliferative disorder,
PT asthma, allergies or AIDS.
XX Claim 9; SEQ ID NO 692; 3205pp; English.
XX The invention relates to novel genes ADA39629-ADA40565 and proteins
CC ADA40566-ADA41501 for human secreted proteins, useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The polypeptides, nucleic acid molecules, antibodies or their
CC fragments, and agonists or antagonists that bind to the polypeptide are
CC useful for preparing a diagnostic or pharmaceutical composition for
CC diagnosing or treating cancer or other hyperproliferative disorder. The
CC polypeptides and nucleic acid molecules are also useful for detecting,
CC preventing, diagnosing, prognosticating, treating or ameliorating cancer
CC or other hyperproliferative disorders including neoplasms, autoimmune
CC disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus
CC erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic
CC anaemia), haematopoietic or haematological disorders (e.g. anaemia,
CC thrombocytopenia), allergic reactions including asthma or eczema,

CC inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory
CC bowel disease or Crohn's disease), neurodegenerative disorders (e.g.
CC Alzheimer's disease or Parkinson's disease), cardiovascular disorders
CC (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,
CC fungal or viral infections including HIV/AIDS), or wound healing and
CC disorders of epithelial cell proliferation. The nucleic acids are also
CC useful for chromosome identification, radiation hybrid mapping or long-
CC range restriction mapping, as molecular weight markers, or as
CC hybridization or diagnostic probes. The polypeptides and antibodies are
CC useful for providing immunological probes for differential identification
CC of the tissues immunohistochemistry assays. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 427 BP; 115 A; 120 C; 102 G; 90 T; 0 U; 0 Other;

Query Match 98.4%; Score 101.4; DB 8; Length 427;
Best Local Similarity 99.0%; Pred. No. 3.3e-27;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTCTCTTCGGTTTATCTGTAGAAATGAA 60

Db 281 CGGACAGCGCAGCAGTCCAGCTCTGGTTCTCTTCGGTTTATCTGTAGAAATGAA 340

Qy 61 ATGGTTCCTCCATAAATAAGGGGCATGAGCCCTTCCTCACAATAA 103

Db 341 ATGGTTCCTCCATAAATAAGGGGCATGAGCCCTTCCTCACAATAA 383

RESULT 7

ACH66744

ID ACH66744 standard; cDNA; 427 BP.

AC ACH66744;

XX 06-NOV-2003 (first entry)

XX Novel human secreted protein #97 cDNA.

XX Human; ss; vaccine; immune system disorder; haematopoietic cell disorder;
KW cancer; autoimmune disorder; rheumatoid arthritis; glomerulonephritis;
KW HIV infection; anaemia; thrombocytopenia blood coagulation disorder;
KW blood platelet disorder; wound; heart attack; myocardial infarction;
KW stroke; scarring; asthma; graft-versus host rejection; inflammation;
KW hyperproliferative disorder; lymphoproliferative disorder; arrhythmia;
KW aberrant cellular division; cell proliferative disorder; angiogenesis;
KW cardiovascular disorder; pulmonary heart disease; neovascularisation;
KW hypertrophic scar; keloid; ocular disorder; diabetic retinopathy; gene;
KW uveitis; epithelial cell proliferation; neurological disease; apoptosis;
KW Parkinson's disease; Alzheimer's disease; Huntington's chorea; AIDS;
KW amyotrophic lateral sclerosis; toxin induced liver disease; septic shock;
KW cachexia; anorexia; lung damage; infection.

XX Homo sapiens.

XX US2003065151-A1.

XX 03-APR-2003.

XX 04-APR-2002; 2002US-00115123.

XX 16-JUN-1998; 98US-0089507P.

XX 16-JUN-1998; 98US-0089508P.

XX 16-JUN-1998; 98US-0089509P.

XX 16-JUN-1998; 98US-0089510P.

XX 22-JUN-1998; 98US-0090112P.

XX 22-JUN-1998; 98US-0090113P.

XX 15-JUN-1999; 99WO-US013418.

XX 14-DEC-1999; 99US-00461325.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Ni J, Rosen CA, Wei Y, Young P, Florence K, Soppet DR;
PI Brewer LA, Endress GA, Carter KC, Mucenski M, Ebner R, Lafleur DW;
XX Olsen H, Shi Y, Moore PA, Komatsoulis G;
DR WPI; 2003-531736/50.
XX P-PSDB; ABO53454.

XX Novel antibody that binds specifically to a HCRJQ69 protein, useful for
PT detecting the presence of a protein in a biological sample, and for
PT treating cancers, autoimmune disorders and HIV infection.

XX Disclosure; SEQ ID NO 107; 176bp; English.

XX The invention relates to an isolated antibody or its fragment that
CC specifically binds to a protein. The antibody is useful for detecting a
CC protein in a biological sample, by contacting the biological sample with
CC the antibody or its fragment and detecting the protein in the biological
CC sample. The antibody is useful for purifying, detecting and targeting the
CC human secreted proteins, including both in vitro and in vivo diagnostic
CC and therapeutic methods. The antibody is useful for immunophenotyping of
CC cell lines in biological samples and in antibody-based therapies for
CC treating, inhibiting and preventing diseases, disorders or conditions
CC associated with aberrant expression and/or activity of the above
CC proteins. The antibody is useful for treating deficiencies or disorders
CC of immune system and haematopoietic cells, for increasing immune
CC and proliferation of haematopoietic cells, for treating immune
CC rheumatoid arthritis and glomerulonephritis), HIV infection, anaemia and
CC thrombocytopenia and as a marker for a particular immune system disease
CC or disorder. The antibody is also useful for treating blood coagulation
CC disorders, blood platelet disorders, wounds, heart attacks (infarction),
CC strokes, scarring and asthma. The antibody is also useful for treating or
CC preventing graft-versus host rejection, for modulating inflammation, for
CC treating hyperproliferative disorders e.g. lymphoproliferative disorders
CC and cancers, for inhibiting aberrant cellular division and for treating
CC cell proliferative disorders. The antibody is also useful for treating
CC cardiovascular disorders e.g. pulmonary heart disease and arrhythmia,
CC disorders associated with neovascularisation and angiogenesis, for
CC treating hypertrophic scars and keloids, ocular disorders e.g. diabetic
CC retinopathy and uveitis, for wound healing and disorders of epithelial
CC cell proliferation. The antibody is also useful for treating neurological
CC diseases e.g. Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea and amyotrophic lateral sclerosis (ALS), diseases associated with
CC increased apoptosis e.g. AIDS, toxin induced liver disease, septic shock,
CC cachexia and anorexia, for preventing and healing damage to lungs and for
CC treating infectious diseases. The present sequence represents cDNA
CC encoding a novel human secreted protein. Note: the sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030065151

XX Sequence 427 BP; 115 A; 120 C; 102 G; 90 T; 0 U; 0 Other;

Query Match 98.4%; Score 101.4; DB 9; Length 427;
Best Local Similarity 99.0%; Pred. No. 3.3e-27;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTCTCTTCGGTTTATCTGTAGAAATGAA 60

Db 281 CGGACAGCGCAGCAGTCCAGCTCTGGTTCTCTTCGGTTTATCTGTAGAAATGAA 340

Qy 61 ATGGTTCCTCCATAAATAAGGGGCATGAGCCCTTCCTCACAATAA 103

Db 341 ATGGTTCCTCCATAAATAAGGGGCATGAGCCCTTCCTCACAATAA 383

RESULT 8

ADB91346

ID ADB91346 standard; cDNA; 427 BP.

XX ADB91346;

XX 04-DEC-2003 (first entry)

```

XX DE Human secreted protein cDNA #SEQ ID 292.
XX KW Secreted protein; gene therapy; antidiabetic; diabetes; human; ss.
XX OS Homo sapiens.
XX PN WO2003004622-A2.
XX PD 16-JAN-2003.
XX XX
XX XX 19-MAR-2002; 2002WO-US008124.
XX XX 21-MAR-2001; 2001US-0277340P.
XX PR 19-JUL-2001; 2001US-0306171P.
XX PR 13-NOV-2001; 2001US-0331287P.
XX XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX XX
XX PI Rosen CA, Ruben SM;
XX DR WPI; 2003-229407/22.
XX XX
XX PT Nucleic acid encoding a human secreted protein is useful in diagnosing or
XX PT treating diabetes or conditions related to diabetes.
XX XX
XX PS Claim 9; SEQ ID NO 292; 1537pp; English.
XX XX
XX CC The invention relates to isolated nucleic acid molecules ADB91065-
XX CC ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-
XX CC ADB91834. Also disclosed is a recombinant vector comprising a
XX CC polynucleotide of the invention, and a recombinant host cell comprising
XX CC the recombinant vector. The polypeptide of the invention is useful in
XX CC identifying a binding partner by contacting the polypeptide with a
XX CC binding partner, and determining whether the binding partner increases or
XX CC decreases activity of the polypeptide. The polypeptide, polynucleotide,
XX CC antibody or its fragment, agonist or antagonist are useful for preparing
XX CC a pharmaceutical composition for diagnosing or treating diabetes or
XX CC conditions related to diabetes. The present sequence is that of the human
XX CC immunoglobulin Fc portion used to generate fusion proteins, increasing
XX CC the stability of the fused protein as compared to the secreted protein
XX CC only. Note: The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 427 BP; 115 A; 120 C; 102 G; 90 T; 0 U; 0 Other;

Query Match 98.4%; Score 101.4; DB 9; Length 427;
Best Local Similarity 99.0%; Pred. No. 3.3e-27;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACAGCGCAGCAGTCAGCTCGGTTCTTCGCTTTATTCGTTAGATGAA 60
Db 281 CGGACAGCGCAGCAGTCAGCTCGGTTCTTCGCTTTATTCGTTAGATGAA 340

Qy 61 ATGGTTCCTCCATTAATAGGGGCATGACCTTCTCTCAAAA 103
Db 341 ATGGTTCCTCCATTAATAGGGGCATGACCTTCTCTCAAAA 383

RESULT 9
AAD34144
ID AAD34144 standard; cDNA; 439 BP.
XX AC
XX AC AAD34144;
XX XX
XX DT 16-JUL-2002 (first entry)
XX DE Human secreted protein-encoding gene 17 cDNA clone HNH023, SEQ ID NO:49.
XX KW Human; secreted protein; proliferative disorder; cancer; tumour; AIDS;
XX KW foetal abnormality; developmental; haematopoietic disorder; kidney; skin;
XX KW immune system; cardiovascular; angiogenic; gastrointestinal; endocrine;

```

```

KW pregnancy-related disorder; cognitive; neurological; autoimmune disease;
KW inflammation; allergy; rheumatoid arthritis; Alzheimer's disease; asthma;
KW Parkinson's disease; schizophrenia; psoriasis; sepsis; atherosclerosis;
KW diabetes; infection; wound healing; vulnery; chemotaxis; food additive;
KW gene therapy; gene; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 12..245
XX FT /*tag= a
XX FT /product= "Human secreted protein"
XX FT /note= "CDS does not include start codon"
XX FT /partial
XX FT sig_peptide 12..71
XX FT /*tag= b
XX FT mat_peptide 72..242
XX FT /*tag= c
XX FT /product= "Mature secreted protein"
XX PN WO200222654-A1.
XX XX
XX PD 21-MAR-2002.
XX XX
XX PF 17-JAN-2001; 2001WO-US001385.
XX XX
XX PR 12-SEP-2000; 2000US-0231969P.
XX XX
XX XX (HUMA-) HUMAN GENOME SCI INC.
XX XX
XX PI Rosen CA, Komatsu GA, Baker KP, Birse CE, Soppet DR;
XX PI Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;
XX PI Fiscella M, Ni J;
XX DR WPI; 2002-315684/35.
XX DR P-PSDB; AAE21616.
XX XX
XX XX Seventeen nucleic acid molecules encoding human secreted proteins, useful
XX XX for treating and preventing cancer, immune disorders (e.g. Addison's
XX XX disease, and allergies), and cardiovascular disorders (e.g. myocardial
XX XX ischemias).
XX PS Claim 1; Page 436-437; 483pp; English.
XX XX
XX CC AAD34106-AAD34144 represent cDNAs corresponding to 17 human secreted
XX CC protein genes, and AAE21578-AAE21616 represent the proteins they encode.
XX CC AAE21617-AAE21627 represent human secreted protein fragments. The genes
XX CC and their corresponding secreted proteins are useful for preventing,
XX CC treating or ameliorating medical conditions, e.g., by protein or gene
XX CC therapy. Pathological conditions can be diagnosed by determining the
XX CC amount of the new protein in a sample or by determining the presence of
XX CC mutations in the new genes. Specific uses are described for each of the
XX CC 17 genes, based on the tissues in which they are most highly expressed,
XX CC and include developing products for the diagnosis or treatment of cancer,
XX CC proliferative disorders, tumours, foetal and developmental abnormalities,
XX CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
XX CC diseases (rheumatoid arthritis), inflammation, allergies, neurological
XX CC disorders (Alzheimer's disease, Parkinson's disease), cognitive disorder,
XX CC schizophrenia, asthma, skin disorders (psoriasis), sepsis, diabetes,
XX CC atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney
XX CC disorders, gastrointestinal disorders, pregnancy-related disorders,
XX CC endocrine disorders, and infections. The proteins can also be used to aid
XX CC wound healing and epithelial cell proliferation, to prevent skin aging
XX CC due to sunburn, to maintain organs before transplantation, for supporting
XX CC cell culture of primary tissues, to regenerate tissues, to identify their
XX CC cognate ligands or binding partners, and in chemotaxis, and can be used
XX CC as a food additive or preservative to modify storage properties. The
XX CC present sequence represents a human secreted protein-encoding cDNA of the
XX CC invention
XX SQ Sequence 439 BP; 124 A; 121 C; 102 G; 90 T; 0 U; 2 Other;

Query Match 98.4%; Score 101.4; DB 6; Length 439;

```

Best Local Similarity		99.0%;	Pred. No. 3.3e-27;	
Matches 102;		Conservative	0;	Mismatches 1;
Indels		0;	Gaps	0;
Qy	1	CGGACAGACGGCAGCAGCTCCAGCTCTGGTTTCTCTTCGGTTTATCTCTGTTAGATGAA	60	
Db	281	CGGACAGACGGCAGCAGCTCCAGCTCTGGTTTCTCTTCGGTTTATCTCTGTTAGATGAA	340	
Qy	61	ATGTTTCCCATTAATAAGGGGCATGAGCCCTTCTCTCAAAAA	103	
Db	341	ATGTTTCCCATTAATAAGGGGCATGAGCCCTTCTCTCAAAAAA	383	
RESULT 10				
ID	AAK56586/c			
ID	AAK56586 standard; cDNA; 1148 BP.			
XX	AC	AAK56586;		
XX	DT	06-NOV-2001 (first entry)		
XX	DE	Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:1646.		
XX	KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;		
XX	KW	cytostatic; gene therapy; vaccine; metastasis; ss.		
XX	OS	Homo sapiens.		
XX	XX	WO200157182-A2.		
XX	PN	09-AUG-2001.		
XX	PD	17-JAN-2001; 2001WO-US001354.		
XX	PF	31-JAN-2000; 2000US-0179065P.		
XX	PR	04-FEB-2000; 2000US-0180628P.		
XX	PR	24-FEB-2000; 2000US-0184664P.		
XX	PR	02-MAR-2000; 2000US-0186350P.		
XX	PR	16-MAR-2000; 2000US-0189874P.		
XX	PR	17-MAR-2000; 2000US-0190076P.		
XX	PR	18-APR-2000; 2000US-0198123P.		
XX	PR	19-MAY-2000; 2000US-0205515P.		
XX	PR	07-JUN-2000; 2000US-0209467P.		
XX	PR	28-JUN-2000; 2000US-0214886P.		
XX	PR	30-JUN-2000; 2000US-0215135P.		
XX	PR	07-JUL-2000; 2000US-0216647P.		
XX	PR	07-JUL-2000; 2000US-0216880P.		
XX	PR	11-JUL-2000; 2000US-0217487P.		
XX	PR	11-JUL-2000; 2000US-0217496P.		
XX	PR	14-JUL-2000; 2000US-0218290P.		
XX	PR	26-JUL-2000; 2000US-0220963P.		
XX	PR	26-JUL-2000; 2000US-0220964P.		
XX	PR	14-AUG-2000; 2000US-0224518P.		
XX	PR	14-AUG-2000; 2000US-0224519P.		
XX	PR	14-AUG-2000; 2000US-0225213P.		
XX	PR	14-AUG-2000; 2000US-0225214P.		
XX	PR	14-AUG-2000; 2000US-0225266P.		
XX	PR	14-AUG-2000; 2000US-0225267P.		
XX	PR	14-AUG-2000; 2000US-0225268P.		
XX	PR	14-AUG-2000; 2000US-0225270P.		
XX	PR	14-AUG-2000; 2000US-0225477P.		
XX	PR	14-AUG-2000; 2000US-0225575P.		
XX	PR	14-AUG-2000; 2000US-0225758P.		
XX	PR	18-AUG-2000; 2000US-0225759P.		
XX	PR	22-AUG-2000; 2000US-0226279P.		
XX	PR	22-AUG-2000; 2000US-0226681P.		
XX	PR	22-AUG-2000; 2000US-0226688P.		
XX	PR	23-AUG-2000; 2000US-0227182P.		
XX	PR	23-AUG-2000; 2000US-0227009P.		
XX	PR	30-AUG-2000; 2000US-0228924P.		
XX	PR	01-SEP-2000; 2000US-0229287P.		
XX	PR	01-SEP-2000; 2000US-0229343P.		
XX	PR	01-SEP-2000; 2000US-0229344P.		
XX	PR	01-SEP-2000; 2000US-0229345P.		

```
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249220P.
PR 17-NOV-2000; 2000US-0249221P.
PR 17-NOV-2000; 2000US-0249222P.
PR 17-NOV-2000; 2000US-0249223P.
PR 17-NOV-2000; 2000US-0249224P.
PR 17-NOV-2000; 2000US-0249225P.
PR 17-NOV-2000; 2000US-0249226P.
PR 17-NOV-2000; 2000US-0249227P.
PR 17-NOV-2000; 2000US-0249228P.
PR 17-NOV-2000; 2000US-0249229P.
PR 17-NOV-2000; 2000US-0249230P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250161P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX P-PSDB; AAM83805.
XX
XX Nucleic acids encoding human immune/haematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Claim 1; SEQ ID NO 1646; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention
XX
XX Sequence 1148 BP; 265 A; 301 C; 364 G; 216 T; 0 U; 2 Other;
XX
XX Query Match 98.4%; Score 101.4; DB 4; Length 1148;
XX Best Local Similarity 99.0%; Pred. No. 4.9e-27;
XX Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 CGGACAGCGGACAGCTCCGAGCTCTGGTTCTCTCGGTTTATCTCTGTTAGATGAA 60
XX
XX 113 CGGACAGCGGACAGCTCCGAGCTCTGGTTCTCTCGGTTTATCTCTGTTAGATGAA 54
XX
XX 61 ATGTTTCCCAATAAAGGGGCGATGAGCCCTTCTCCACAAAAA 103
XX
XX 53 ATGTTTCCCAATAAAGGGGCGATGAGCCCTTCTCCACAAAAA 11
XX
XX
XX RESULT 11
XX ADI62940
XX ID ADI62940 standard; cDNA; 1874 BP.
XX
AC ADI62940;
XX
XX 22-APR-2004 (first entry)
XX
XX Human apoptosis-associated cDNA SEQ ID 383.
XX
XX apoptosis; cell death; cytostatic; neuroprotective; immunosuppressive;
XX antirheumatic; antiarthritic; dermatological; antiinflammatory;
XX hepatotropic; virucide; nontropic; anticonvulsant; antiparkinsonian;
XX vasotropic; cerebroprotective; antialcoholic; gene therapy; tumour;
XX autoimmune disease; degenerative disease; viral infection; leukaemia;
XX carcinoma; sarcoma; multiple sclerosis; rheumatoid arthritis; diabetes;
XX lupus; hepatitis; influenza viruses; Alzheimer's disease;
XX Huntington's disease; Parkinson's diseases; reperfusion injury; stroke;
XX alcoholic liver disease; human; gene; ss.
XX
XX Homo sapiens.
XX
XX OS
XX WO2003058021-A2.
XX
XX PN
XX 17-JUL-2003.
XX
XX PD
XX
XX PF 13-JAN-2003; 2003WO-EP000270.
XX
XX PR 11-JAN-2002; 2002DE-01000856.
XX
XX PR (XANT-) XANTOS BIOMEDICINE AG.
XX
XX Koenig-Hoffman K, Kazinski M, Schaefer R, Kesper B;
XX WPI; 2003-542134/51.
XX
XX
XX New nucleic acids involved in apoptosis, useful for diagnosis and
XX treatment of e.g. tumors and degenerative disease, also related proteins,
XX antibodies and modulators.
XX
XX Claim 1b; SEQ ID NO 383; 517pp; German.
XX
XX This invention describes novel nucleic acid molecules that are associated
XX with apoptosis and encode a polypeptide and are derived from a normalised
XX gene library (embryonic or liver) or clone collections, and the extent of
XX apoptosis measured by cell death detection assay or the CPBG assay
XX (measuring loss of membrane integrity). The products of the invention
XX have cytostatic, neuroprotective, immunosuppressive, antirheumatic,
XX antiarthritic, dermatological, antiinflammatory, hepatotropic, virucide,
XX nontropic, anticonvulsant, antiparkinsonian, vasotropic,
XX cerebroprotective and antialcoholic activity and can be used for gene
XX therapy. The polynucleotides also related vectors, hosts (or their
XX extracts), encoded polypeptide (or their receptors) and/or agents that
XX inhibit their activity (including antisense sequences) are used for
XX treatment or prevention of tumours, autoimmune or degenerative diseases
XX and viral infections, specifically leukaemia, carcinoma, sarcoma,
XX multiple sclerosis, rheumatoid arthritis, diabetes, lupus, or infection
XX with hepatitis or influenza viruses, Alzheimer's, Huntington's or
XX Parkinson's diseases, reperfusion injury, stroke and alcoholic liver
XX disease. Detection of the polynucleotides and derived polypeptides can
XX also be used for diagnosis of these diseases. This sequence encodes an
XX apoptosis-associated protein described in the disclosure of the
XX invention.
XX
XX Sequence 1874 BP; 343 A; 632 C; 546 G; 353 T; 0 U; 0 Other;
XX
XX Query Match 98.4%; Score 101.4; DB 10; Length 1874;
XX Best Local Similarity 99.0%; Pred. No. 6e-27;
XX Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 CGGACAGCGGACAGCTCCGAGCTCTGGTTCTCTCGGTTTATCTCTGTTAGATGAA 60
XX
XX 1707 CGGACAGCGGACAGCTCCGAGCTCTGGTTCTCTCGGTTTATCTCTGTTAGATGAA 1766
XX
XX 61 ATGTTTCCCAATAAAGGGGCGATGAGCCCTTCTCCACAAAAA 103
XX
XX 1767 ATGTTTCCCAATAAAGGGGCGATGAGCCCTTCTCCACAAAAA 1809
XX
```

thrombolytic; anticoagulant; antilipaeamic; hypotensive; cardiant; peptide/histidine transporter 2; PHT2.

```
RESULT 12
AAI68202
ID AAI68202 standard; cDNA; 2031 BP.
XX
XX AC AAI68202;
XX
XX DT 18-DEC-2001 (first entry)
XX
XX DE Human osteoclast expressed polynucleotide SEQ ID NO 5.
XX
XX KW Human; osteoclast; ss.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX CDS 128..1873
XX FT /*tag= a
XX FT /product= "osteoclast expressed protein"
XX
XX JP2001231573-A.
XX
XX PD 28-AUG-2001.
XX
XX PF 21-FEB-2000; 2000JP-00048886.
XX
XX PR 21-FEB-2000; 2000JP-00048886.
XX
XX PA (SHIO ) SHIONOGI & CO LTD.
XX
XX WPI; 2001-609964/70.
XX DR P-PSDB; AAM51210.
XX
XX PT Gene derived from human osteoclast, used as an index for the maturation
XX of osteoclasts.
XX
XX PS Claim 1; Page 20-22; 31pp; Japanese.
XX
XX CC The invention relates to human osteoclast expressed polynucleotide
XX sequences (AAI68200-AAI68203) and the encoded proteins (AAM51208-
XX AAM51211). The polynucleotides are useful as an index for the maturation
XX of osteoclasts
XX
XX SQ Sequence 2031 BP; 327 A; 684 C; 611 G; 409 T; 0 U; 0 Other;

Query Match 98.4%; Score 101.4; DB 5; Length 2031;
Best Local Similarity 99.0%; Pred. No. 6.2e-27;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGACGGCAGCAGTCCAGCTCGGTTCTCTCGGTTTATTCTGTAGATGAA 60
DB 1909 CGGACAGACGGCAGCAGTCCAGCTCGGTTCTCTCGGTTTATTCTGTAGATGAA 1968

QY 61 ATGGTTCCCAATAAATAGGGGCATGAGCCCTTCCTCACAAAAA 103
DB 1969 ATGGTTCCCAATAAATAGGGGCATGAGCCCTTCCTCACAAAAA 2011

RESULT 13
ADQ88225
ID ADQ88225 standard; DNA; 2057 BP.
XX
XX AC ADQ88225;
XX
XX DT 21-OCT-2004 (first entry)
XX
XX DE Human 33556 DNA encodes a peptide/histidine transporter 2 protein PHT2.
XX
XX KW human; gene; ds; cardiovascular disorder; thrombotic disorder;
XX differential expression; gene therapy; aberrant vascularisation;
XX atherosclerosis; thrombosis; coronary artery disease; hyperlipidaemia;
XX dyslipidaemia; high blood pressure; heart failure; cardiant;

KW KW thrombolytic; anticoagulant; antilipaeamic; hypotensive; cardiant;
KW peptide/histidine transporter 2; PHT2.
OS OS Homo sapiens.
XX WO2004063340-A2.
XX
XX PD 29-JUL-2004.
XX
XX PF 13-JAN-2004; 2004WO-US000393.
XX
XX PR 13-JAN-2003; 2003US-0439683P.
XX PR 05-FEB-2003; 2003US-0445216P.
XX PR 18-FEB-2003; 2003US-0448036P.
XX PR 12-MAR-2003; 2003US-0454189P.
XX PR 25-MAR-2003; 2003US-0457541P.
XX PR 29-APR-2003; 2003US-0466411P.
XX PR 08-MAY-2003; 2003US-0469041P.
XX PR 10-JUN-2003; 2003US-0477414P.
XX PR 13-JUN-2003; 2003US-0478560P.
XX PR 24-JUL-2003; 2003US-0489772P.
XX PR 28-JUL-2003; 2003US-0490660P.
XX PR 03-SEP-2003; 2003US-0499838P.
XX PR 22-SEP-2003; 2003US-0504786P.
XX PR 24-SEP-2003; 2003US-050570P.
XX PR 17-OCT-2003; 2003US-0512418P.
XX PR 27-OCT-2003; 2003US-0514660P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX STagliano NE, Healy A, Acton SL, Galvin KM, Donoghue MA;
XX Rogriague-Way A, Tomlinson JE;
XX WPI; 2004-553729/53.
XX DR P-PSDB; ADQ88226.
XX
XX PT Identifying a compound for treating a cardiovascular or thrombotic
XX disorder by combining a compound to be tested with e.g., a 9380, 9462,
XX 8701 or 2419 polypeptide or with a host cell expressing the polypeptide
XX and detecting the binding.
XX
XX PS Claim 1; SEQ ID NO 67; 512pp; English.
XX
XX CC This invention relates to a novel compound that is capable of treating a
XX cardiovascular or thrombotic disorder. Specifically, it refers to the
XX identification of nucleic acid molecules, and the encoded proteins
XX thereof, which are differentially expressed in cardiovascular disease
XX states relative to their normal expression in non-diseased tissue. The
XX present invention describes test compounds (i.e. small molecules,
XX peptides or antibodies) that can bind to and modulate the activity of
XX these differentially expressed membrane-bound polypeptides, where binding
XX is detected by a competition binding assay, immunoassay or yeast two-
XX hybrid assay. Accordingly, pharmaceutical compositions can be developed
XX and used via gene therapy to treat aberrant vascularisation,
XX atherosclerosis, thrombosis, coronary artery disease, hyperlipidaemia,
XX dyslipidaemia, high blood pressure or heart failure. As such, they
XX exhibit cardiant, thrombolytic, anticoagulant, antilipaeamic, hypotensive
XX and cardiant activities. This polynucleotide sequence is a human DNA
XX molecule that is differentially expressed in a patient with a
XX cardiovascular disorder, given in an exemplification of the invention.
XX
XX SQ Sequence 2057 BP; 331 A; 692 C; 619 G; 415 T; 0 U; 0 Other;

Query Match 98.4%; Score 101.4; DB 13; Length 2057;
Best Local Similarity 99.0%; Pred. No. 6.2e-27;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGACGGCAGCAGTCCAGCTCGGTTCTCTCGGTTTATTCTGTAGATGAA 60
DB 1914 CGGACAGACGGCAGCAGTCCAGCTCGGTTCTCTCGGTTTATTCTGTAGATGAA 1973

QY 61 ATGGTTCCCAATAAATAGGGGCATGAGCCCTTCCTCACAAAAA 103
|||||
```

Db 1974 ATGGTCCCATAAATAAGGGCGATGAGCCCTTCTCAGCAAAA 2016

RESULT 14

AAI70679

ID AAI70679 standard; cDNA; 2060 BP.

AC AAI70679;

XX

XX 04-FEB-2002 (first entry)

DT Human membrane transporter (MTP) 33556 cDNA clone Fbh33556.

XX

XX Membrane transporter; MTP; human; cell proliferation; cell growth;

KW cell differentiation; cell migration; immune response; hormone response;

KW cell communication; nootropic; neuroprotective; anticonvulsant;

KW hypotensive; antidepressant; neuroleptic; antiarteriosclerotic;

KW vasotropic; antianginal; cardiant; thrombolytic; cytostatic;

KW antiinfertility; immunosuppressive; immunostimulant;

KW proton-dependent oligopeptide transport; POT; diagnosis; gene therapy;

XX ss.

XX Homo sapiens.

OS

XX

XX Key Location/Qualifiers

PH 136..1746

FT /*cag= a

FT /notes= "the coding region is also claimed in Claim 1(a)"

FT

XX W0200179497-A2.

XX

XX 25-OCT-2001.

XX

XX 13-APR-2001; 2001WO-US012187.

XX

XX 14-APR-2000; 2000US-0197376P.

XX

XX (MTLL-) MILLENNIUM PHARM INC.

XX

XX Curtis RAJ;

XX

XX WPI; 2002-034357/04.

XX

XX P-PSDB; AAM50318.

XX

XX Novel 33556 (a membrane transporter) protein and nucleic acids encoding

PI the protein, useful for treating disorders associated with transporter-

PI mediated activity such as Alzheimer's disease, arteriosclerosis, cancer.

XX

XX Claim 1(a); Fig 1A-D; 115pp; English.

XX

XX The present sequence, the coding region of which is also claimed, is that

CC of cDNA clone Fbh33556, which codes for novel 581-amino acid (64 kDa)

CC human membrane transporter (MTP) protein 33556 (see AAM50318). MTP is a

CC novel member of the proton-dependent oligopeptide transport (POT) family

CC of proteins possessing the ability to shuttle molecules across a lipid

CC bilayer. MTP proteins are capable of transporting ions, proteins, and

CC small molecules across biological membranes both within a cell and

CC between the cell and the environment and, thus, play a role in or

CC function in a variety of cellular processes, e.g. proliferation, growth,

CC differentiation, migration, immune responses, hormonal responses, and

CC inter- or intracellular communication. The invention provides MTP nucleic

CC acids, antisense nucleic acids, recombinant expression vectors, host

CC cells and nonhuman transgenic animals in which an MTP gene has been

CC introduced or disrupted, as well as MTP proteins, antigenic peptides and

CC anti-MTP antibodies. These are used in claimed methods for identifying a

CC subject at risk of developing an MTP disorder, and for identifying a

CC modulator compounds capable of treating such as disorder. The 33556

CC modulator may be an antisense 33556 nucleic acid, a ribozyme, the present

CC nucleic acid sequence or its fragment or variant. 33556 Nucleic acids are

CC also used in the recombinant production of 33556 polypeptides, and as

CC probes and primers. Examples of MTP-associated or related disorders

CC include: transporter-associated disorders such as CNS disorders including

CC Alzheimer's disease, senile dementia, Huntington's disease, multiple

CC

CC sclerosis, amyotrophic lateral sclerosis, Jakob-Creutzfeldt disease,

CC autonomic function disorders such as hypertension and sleep disorders,

CC and neuropsychiatric disorders such as depression, schizophrenia,

CC learning or memory disorders such as amnesia or age-related memory loss,

CC attention deficit disorder, dysthymic disorder, depression, mania,

CC obsessive compulsive disorder, anxiety, phobia, panic, and bipolar

CC affective disorder; cardiac-related disorders such as arteriosclerosis,

CC ischaemic reperfusion injury, restenosis, arterial inflammation, vascular

CC wall remodeling, coronary microembolism, tachycardia, bradycardia,

CC anginal, hypertension, idiopathic cardiomyopathy, myocardial infarction;

CC cancer such as carcinoma, sarcoma or leukaemia, metastasis, skeletal

CC dysplasia, hepatic disorders and haematopoietic and/or myeloproliferative

CC disorders; hormonal disorders such as type I and type II diabetes,

CC pituitary disorders, thyroid disorders, reproductive or fertility

CC disorders; autoimmune or immune deficiency disorders, such as congenital

CC X-linked infantile hypogammaglobulinaemia, transient

CC hypogammaglobulinaemia, selective IGA deficiency; and muscular disorders

CC such as paralysis, muscle weakness, muscular dystrophy, congenital

CC myopathies. The 33556 molecules are also useful as markers of disorders

CC or disease states, as markers for precursors of disease states, as

CC markers for predisposition of disease states, as markers of drug

CC activity, or as markers of the pharmacogenomic profile of a subject

XX

SQ Sequence 2060 BP; 330 A; 694 C; 613 G; 415 T; 0 U; 8 Other;

Query Match 98.4%; Score 101.4; DB 6; Length 2060;

Best Local Similarity 99.0%; Pred. No. 6.2e-27;

Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATTCTGTAGATGAA 60

DB 1917 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATTCTGTAGATGAA 1976

QY 61 ATGGTTCCCATAAATAAGGGCGATGAGCCCTTCTCAGCAAAA 103

DB 1977 ATGGTTCCCATAAATAAGGGCGATGAGCCCTTCTCAGCAAAA 2019

RESULT 15

AAK52347

ID AAK52347 standard; cDNA; 2124 BP.

XX

XX AAK52347;

XX

XX 06-NOV-2001 (first entry)

DT Human polynucleotide SEQ ID NO 892.

XX

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorder; arthritis; inflammation; ss.

XX

XX Homo sapiens.

XX

XX W0200157190-A2.

XX

XX 09-AUG-2001.

XX

XX 05-FEB-2001; 2001WO-US0004098.

XX

XX 03-FEB-2000; 2000US-00496914.

XX

XX 27-APR-2000; 2000US-00560875.

XX

XX 20-JUN-2000; 2000US-00598075.

XX

XX 19-JUL-2000; 2000US-00620325.

XX

XX 01-SEP-2000; 2000US-00654936.

XX

XX 15-SEP-2000; 2000US-00663561.

XX

XX 20-OCT-2000; 2000US-00693325.

XX

XX 30-NOV-2000; 2000US-00728422.

XX

XX (HYSE-) HYSEQ INC.

XX

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

PI

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:31:40 ; Search time 256.473 Seconds
(without alignments)
3320.997 Million cell updates/sec

Title: US-09-980-046B-8

Perfect score: 103

Sequence: 1 cggacagacggcagcagtc.....tgagcccttcctcacaaaa 103

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA_Main:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102.2	99.2	1149	6	US-10-264-237-500
2	101.4	98.4	427	5	US-10-012-542-107
3	101.4	98.4	427	5	US-10-115-123-107
4	101.4	98.4	427	7	US-10-800-834-107
5	101.4	98.4	1874	8	US-10-775-920-276
6	101.4	98.4	2034	8	US-10-775-920-277
7	101.4	98.4	2057	3	US-09-835-270-1
8	101.4	98.4	2057	3	US-10-753-267-67
9	100.4	97.5	1277	5	US-10-012-542-20
10	100.4	97.5	1277	5	US-10-115-123-20
11	100.4	97.5	1277	7	US-10-800-834-20
12	98.4	95.5	26555	3	US-09-860-670-161
13	98.4	95.5	26555	6	US-10-227-646-161
14	98.4	95.5	136328	6	US-10-101-510-127
15	96.4	93.6	2113	6	US-10-172-118-1657
16	96.4	93.6	2113	7	US-10-342-887-1657
17	96.4	93.6	2113	8	US-10-755-889-203
18	96.4	93.6	2113	8	US-10-775-920-274
19	96.4	93.6	2113	8	US-10-775-920-275
20	90.4	87.8	1104	7	US-10-276-774-720
21	74.4	72.2	280	3	US-09-745-288-58
22	74.4	72.2	280	7	US-10-453-919-58
23	74.4	72.2	2786	3	US-09-814-353-20133

c	24	72.8	70.7	2755	6	US-10-269-909-32	Sequence 32, Appl
c	25	72.8	70.7	2755	8	US-10-723-860-1705	Sequence 1705, Ap
c	26	72.8	70.7	2755	9	US-10-936-626-33	Sequence 33, Appl
c	27	72.8	70.7	2755	9	US-10-938-061-33	Sequence 33, Appl
c	28	72.8	70.7	2769	8	US-10-723-860-6089	Sequence 6089, Ap
c	29	71.8	69.7	345	6	US-10-062-674-1501	Sequence 1501, Ap
c	30	70	68.0	70	9	US-10-957-432-406	Sequence 406, App
c	31	68	66.0	1902	6	US-10-257-022-26	Sequence 26, Appl
c	32	63.8	61.9	317	3	US-09-968-007A-622	Sequence 622, App
c	33	63.8	61.9	317	3	US-09-968-007A-914	Sequence 914, App
c	34	63.8	61.9	317	9	US-10-843-641A-7092	Sequence 7092, Ap
c	35	63.8	61.9	317	9	US-10-843-641A-7384	Sequence 7384, Ap
c	36	43	41.7	281	3	US-09-960-352-1755	Sequence 1755, Ap
c	37	40.2	39.0	346	6	US-10-062-674-362	Sequence 362, App
c	38	33.4	32.4	1132	4	US-09-925-065A-92388	Sequence 92388, A
c	39	32.2	31.3	594	3	US-09-764-891-536	Sequence 536, App
c	40	31	30.1	1401	7	US-10-282-122A-17398	Sequence 17398, A
c	41	28.8	28.0	222	3	US-09-783-590-636	Sequence 636, App
c	42	28.8	28.0	553	4	US-09-925-065A-517586	Sequence 517586, A
c	43	28.8	28.0	1284	7	US-10-425-114-10355	Sequence 10355, A
c	44	28.8	28.0	2366	7	US-10-424-599-39215	Sequence 39215, A
c	45	28.8	28.0	2672	7	US-10-424-599-39212	Sequence 39212, A

ALIGNMENTS

RESULT 1
US-10-264-237-500
; Sequence 500, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birste et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 500
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (99)..(99)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1133)..(1134)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1144)..(1144)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1149)..(1149)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-500

Query Match 99.2%; Score 102.2; DB 6; Length 1149;
Best Local Similarity 98.1%; Pred. No. 1.9e-26;
Matches 101; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGACAGCGCAGCAGTCCTCGGTTTCCTCGGTTTATTCTGTAGATGAA 60
DB 991 CGGACAGCGCAGCAGTCCTCGGTTTCCTCGGTTTATTCTGTAGATGAA 1050

QY 61 ATGTTCCCATTAATAAGGGCATGAGCCCTTCTCACAATAA 103
|||||
Db 1051 ATGTTCCCATTAATAAGGGCATGAGCCCTTCTCACAATAA 1093
|||||

RESULT 2

US-10-012-542-107
; Sequence 107, Application US/10012542
; Publication No. US20030044851A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; SOFTWARE: PatentIn Ver. 2.0
; NUMBER OF SEQ ID NOS: 532
; SEQ ID NO 107
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-542-107

Query Match 98.4%; Score 101.4; DB 5; Length 427;
Best Local Similarity 99.0%; Pred. No. 2.5e-26;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATCTCTGTTAGATGAA 60
|||||
Db 281 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATCTCTGTTAGATGAA 340
|||||
QY 61 ATGTTCCCATTAATAAGGGCATGAGCCCTTCTCACAATAA 103
|||||
Db 341 ATGTTCCCATTAATAAGGGCATGAGCCCTTCTCACAATAA 383
|||||

RESULT 3

US-10-115-123-107
; Sequence 107, Application US/10115123
; Publication No. US20030065151A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029G3OAP1D2
; CURRENT APPLICATION NUMBER: US/10/115,123
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: PCT/US99/13418
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113

; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 107
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-115-123-107

Query Match 98.4%; Score 101.4; DB 5; Length 427;
Best Local Similarity 99.0%; Pred. No. 2.5e-26;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATCTCTGTTAGATGAA 60
|||||
Db 281 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATCTCTGTTAGATGAA 340
|||||
QY 61 ATGTTCCCATTAATAAGGGCATGAGCCCTTCTCACAATAA 103
|||||
Db 341 ATGTTCCCATTAATAAGGGCATGAGCCCTTCTCACAATAA 383
|||||

RESULT 4

US-10-800-834-107
; Sequence 107, Application US/10800834
; Publication No. US20040146930A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1D3
; CURRENT APPLICATION NUMBER: US/10/800,834
; CURRENT FILING DATE: 2004-03-16
; PRIOR APPLICATION NUMBER: 10/115,123
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: PCT/US99/13418
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 107
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-800-834-107

Query Match 98.4%; Score 101.4; DB 7; Length 427;
Best Local Similarity 99.0%; Pred. No. 2.5e-26;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATCTCTGTTAGATGAA 60
|||||
Db 281 CGGACAGACGGCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATCTCTGTTAGATGAA 340
|||||
QY 61 ATGTTCCCATTAATAAGGGCATGAGCCCTTCTCACAATAA 103
|||||
Db 341 ATGTTCCCATTAATAAGGGCATGAGCCCTTCTCACAATAA 383
|||||

RESULT 5

US-10-775-920-276

```
; Sequence 276, Application US/10775920
; Publication No. US2004017544A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 276
; LENGTH: 1874
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-276

Query Match      98.4%; Score 101.4; DB 8; Length 1874;
Best Local Similarity 99.0%; Pred. No. 4.7e-26;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTCTCGTTTATTCGTTAGAAATGAA 60
DB 1707 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTCTCGTTTATTCGTTAGAAATGAA 1766

QY 61 ATGGTTCCTCCATAATAAGGGGCATGAGCCCTTCCTCACAAAAA 103
DB 1767 ATGGTTCCTCCATAATAAGGGGCATGAGCCCTTCCTCACAAAAA 1809

RESULT 6
US-10-775-920-277
; Sequence 277, Application US/10775920
; Publication No. US2004017544A1
; GENERAL INFORMATION:
; APPLICANT: Mergen Ltd
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES AND CORRESPONDING ENCODED POLYPEPTIDES
; TITLE OF INVENTION: OF PARTICULAR SECRETED AND MEMBRANE-BOUND PROTEINS OVEREXPRESSED
; FILE REFERENCE: Mergen - 0010B
; CURRENT APPLICATION NUMBER: US/10/775,920
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 60/447,900
; PRIOR FILING DATE: 2003-02-13
; NUMBER OF SEQ ID NOS: 385
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 277
; LENGTH: 2034
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-920-277

Query Match      98.4%; Score 101.4; DB 8; Length 2034;
Best Local Similarity 99.0%; Pred. No. 4.7e-26;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTCTCGTTTATTCGTTAGAAATGAA 60
DB 1876 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTCTCGTTTATTCGTTAGAAATGAA 1935

QY 61 ATGGTTCCTCCATAATAAGGGGCATGAGCCCTTCCTCACAAAAA 103
DB 1936 ATGGTTCCTCCATAATAAGGGGCATGAGCCCTTCCTCACAAAAA 1978

RESULT 7
US-09-835-270-1
; Sequence 1, Application US/09835270
; Patent No. US20020077462A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Curtis, Rory A.
; TITLE OF INVENTION: 33556, A NOVEL HUMAN TRANSPORTER AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 381553000800
; CURRENT APPLICATION NUMBER: US/09/835,270
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,376
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2057
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)...(1878)
; OTHER INFORMATION: cDNA of human MTP
US-09-835-270-1

Query Match      98.4%; Score 101.4; DB 3; Length 2057;
Best Local Similarity 99.0%; Pred. No. 4.7e-26;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTCTCGTTTATTCGTTAGAAATGAA 60
DB 1914 CGGACAGCGCAGCAGTCCAGCTCTGGTTTCCTTCGCTTCTCGTTTATTCGTTAGAAATGAA 1973

QY 61 ATGGTTCCTCCATAATAAGGGGCATGAGCCCTTCCTCACAAAAA 103
DB 1974 ATGGTTCCTCCATAATAAGGGGCATGAGCCCTTCCTCACAAAAA 2016

RESULT 8
US-10-753-267-67
; Sequence 67, Application US/10753267
; Publication No. US20050037946A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Stagliano, Nancy E.
; APPLICANT: Healy, Aileen L.
; APPLICANT: Acton, Susan L.
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Donoghue, Mary A.
; APPLICANT: Rodrigue-Way, Amelie
; APPLICANT: Tomlinson, James E.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE USING 1722, 10280, 59917, 85553,
; TITLE OF INVENTION: 10653, 9235, 21668, 17794, 2210, 6169, 10102, 21061, 17662,
; TITLE OF INVENTION: 1468, 12282, 6350, 9035, 1820, 23652, 7301, 8925, 8701,
; TITLE OF INVENTION: 3533, 9462, 9123, 12788, 17729, 65552, 1261, 21476, 33770,
; TITLE OF INVENTION: 9380, 2565654, 33556, 44143, 32612, 10671, 261,
; TITLE OF INVENTION: 44570, 41922, 2552, 2417, 19319, 43969, 8921, 8993, 955,
; TITLE OF INVENTION: 32345, 966, 1920, 17318, 1510, 14180, 26005, 554, 16408,
; TITLE OF INVENTION: 42028, 112091, 13886, 13942, 1673, 54946 OR 2419
; FILE REFERENCE: MPI03-003PIRNONMIN
; CURRENT APPLICATION NUMBER: US/10/753,267
; CURRENT FILING DATE: 2004-01-08
; PRIOR APPLICATION NUMBER: US 60/439,683
; PRIOR FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: US 60/445,216
; PRIOR FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US 60/448,036
; PRIOR FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/454,189
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: US 60/457,541
; PRIOR FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US 60/466,411
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/469,041
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/477,414
```

```
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: US 60/478,560
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/489,772
; PRIOR FILING DATE: 2003-07-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 67
; LENGTH: 2057
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)...(1878)
US-10-753-267-67

Query Match          98.4%; Score 101.4; DB 8; Length 2057;
Best Local Similarity 99.0%; Pred. No. 4.7e-26;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTTAGAA 60
Db      1914 CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTTAGAA 60
QY      61  ATGTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAATAA 103
Db      1974 ATGTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAATAA 2016

RESULT 9
US-10-012-542-20
; Sequence 20, Application US/10012542
; Publication No. US20030044851A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1277
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1207)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1272)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-012-542-20

Query Match          97.5%; Score 100.4; DB 5; Length 1277;
Best Local Similarity 98.1%; Pred. No. 9e-26;
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTTAGAA 60
Db      1914 CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTTAGAA 60
QY      61  ATGTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAATAA 103
Db      1974 ATGTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAATAA 2016

RESULT 10
US-10-115-123-20
; Sequence 20, Application US/10115123
; Publication No. US20030065151A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029G30AP1D2
; CURRENT APPLICATION NUMBER: US/10/115,123
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: PCT/US99/13418
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1277
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1207)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1272)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-115-123-20

Query Match          97.5%; Score 100.4; DB 5; Length 1277;
Best Local Similarity 98.1%; Pred. No. 9e-26;
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTTAGAA 60
Db      1130 CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTTAGAA 1189
QY      61  ATGTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAATAA 103
Db      1190 ATGTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAATAA 1232

RESULT 11
US-10-800-834-20
; Sequence 20, Application US/10800834
; Publication No. US20040146930A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1D3
; CURRENT APPLICATION NUMBER: US/10/800,834
; CURRENT FILING DATE: 2004-03-16
; PRIOR APPLICATION NUMBER: 10/115,123
; PRIOR FILING DATE: 2002-04-04
```

```
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: US 60/478,560
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/489,772
; PRIOR FILING DATE: 2003-07-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 67
; LENGTH: 2057
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)...(1878)
US-10-753-267-67

Query Match          98.4%; Score 101.4; DB 8; Length 2057;
Best Local Similarity 99.0%; Pred. No. 4.7e-26;
Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTTAGAA 60
Db      1914 CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTTAGAA 60
QY      61  ATGTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAATAA 103
Db      1974 ATGTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAATAA 2016

RESULT 9
US-10-012-542-20
; Sequence 20, Application US/10012542
; Publication No. US20030044851A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1277
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1207)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1272)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-012-542-20

Query Match          97.5%; Score 100.4; DB 5; Length 1277;
Best Local Similarity 98.1%; Pred. No. 9e-26;
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTTAGAA 60
Db      1914 CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTTAGAA 60
QY      61  ATGTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAATAA 103
Db      1974 ATGTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAATAA 2016

RESULT 10
US-10-115-123-20
; Sequence 20, Application US/10115123
; Publication No. US20030065151A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029G30AP1D2
; CURRENT APPLICATION NUMBER: US/10/115,123
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: PCT/US99/13418
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1277
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1207)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1272)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-115-123-20

Query Match          97.5%; Score 100.4; DB 5; Length 1277;
Best Local Similarity 98.1%; Pred. No. 9e-26;
Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTTAGAA 60
Db      1130 CGACAGACGGCAGCAGTCCAGCTCTGGTTCTCTCGGTTTATTCGTTAGAA 1189
QY      61  ATGTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAATAA 103
Db      1190 ATGTTCCCAATAAATAGGGGCATGAGCCCTTCTCACAATAA 1232

RESULT 11
US-10-800-834-20
; Sequence 20, Application US/10800834
; Publication No. US20040146930A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1D3
; CURRENT APPLICATION NUMBER: US/10/800,834
; CURRENT FILING DATE: 2004-03-16
; PRIOR APPLICATION NUMBER: 10/115,123
; PRIOR FILING DATE: 2002-04-04
```

; PRIOR APPLICATION NUMBER: 09/461,325
 ; PRIOR FILING DATE: 1999-12-14
 ; PRIOR APPLICATION NUMBER: PCT/US99/13418
 ; PRIOR FILING DATE: 1999-06-15
 ; PRIOR APPLICATION NUMBER: 60/089,507
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089,508
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089,509
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089,510
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/090,112
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090,113
 ; PRIOR FILING DATE: 1998-06-22
 ; NUMBER OF SEQ ID NOS: 532
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 20
 ; LENGTH: 1277
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (1207)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (1272)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; US-10-800-834-20
 Query Match 97.5%; Score 100.4; DB 7; Length 1277;
 Best Local Similarity 98.1%; Pred. No. 9e-26;
 Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CGGACAGACGGCAGCAGTCCAGCTCGTTTCCTCTCGGTTTATTCGTTAGATGAA 60
 DB 1130 CGGACAGACGGCAGCAGTCCAGCTCGTTTCCTCTCGGTTTATTCGTTAGATGAA 1189
 QY 61 ATGGTTCCTCCATAAATAGGGGCATGAGCCCTTCCTCACAATAA 103
 DB 1190 ATGGTTCCTCCATAAATAGGGGCATGAGCCCTTCCTCACAATAA 1232
 RESULT 12
 US-09-860-670-161
 ; Sequence 161, Application US/09860670
 ; Patent No. US20020165137A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PA127P1
 ; CURRENT APPLICATION NUMBER: US/09/860,670
 ; CURRENT FILING DATE: 2001-05-21
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 289
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 161
 ; LENGTH: 26555
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-860-670-161
 Query Match 95.5%; Score 98.4; DB 3; Length 26555;
 Best Local Similarity 99.0%; Pred. No. 1.6e-24;
 Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGGACAGACGGCAGCAGTCCAGCTCGTTTCCTCTCGGTTTATTCGTTAGATGAA 60
 DB 13862 CGGACAGACGGCAGCAGTCCAGCTCGTTTCCTCTCGGTTTATTCGTTAGATGAA 13921
 QY 61 ATGGTTCCTCCATAAATAGGGGCATGAGCCCTTCCTCACAATAA 100
 DB 13759 ATGGTTCCTCCATAAATAGGGGCATGAGCCCTTCCTCACAATAA 37398
 ; PRIOR APPLICATION NUMBER: 09/461,325
 ; PRIOR FILING DATE: 1999-12-14
 ; PRIOR APPLICATION NUMBER: PCT/US99/13418
 ; PRIOR FILING DATE: 1999-06-15
 ; PRIOR APPLICATION NUMBER: 60/089,507
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089,508
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089,509
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/089,510
 ; PRIOR FILING DATE: 1998-06-16
 ; PRIOR APPLICATION NUMBER: 60/090,112
 ; PRIOR FILING DATE: 1998-06-22
 ; PRIOR APPLICATION NUMBER: 60/090,113
 ; PRIOR FILING DATE: 1998-06-22
 ; NUMBER OF SEQ ID NOS: 532
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 20
 ; LENGTH: 1277
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (1207)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (1272)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; US-10-800-834-20
 Query Match 97.5%; Score 100.4; DB 7; Length 1277;
 Best Local Similarity 98.1%; Pred. No. 9e-26;
 Matches 101; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CGGACAGACGGCAGCAGTCCAGCTCGTTTCCTCTCGGTTTATTCGTTAGATGAA 60
 DB 1130 CGGACAGACGGCAGCAGTCCAGCTCGTTTCCTCTCGGTTTATTCGTTAGATGAA 1189
 QY 61 ATGGTTCCTCCATAAATAGGGGCATGAGCCCTTCCTCACAATAA 103
 DB 1190 ATGGTTCCTCCATAAATAGGGGCATGAGCCCTTCCTCACAATAA 1232
 RESULT 12
 US-09-860-670-161
 ; Sequence 161, Application US/09860670
 ; Patent No. US20020165137A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PA127P1
 ; CURRENT APPLICATION NUMBER: US/09/860,670
 ; CURRENT FILING DATE: 2001-05-21
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 289
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 161
 ; LENGTH: 26555
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-860-670-161
 Query Match 95.5%; Score 98.4; DB 3; Length 26555;
 Best Local Similarity 99.0%; Pred. No. 1.6e-24;
 Matches 99; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CGGACAGACGGCAGCAGTCCAGCTCGTTTCCTCTCGGTTTATTCGTTAGATGAA 60
 DB 13862 CGGACAGACGGCAGCAGTCCAGCTCGTTTCCTCTCGGTTTATTCGTTAGATGAA 13921
 QY 61 ATGGTTCCTCCATAAATAGGGGCATGAGCCCTTCCTCACAATAA 100
 DB 13759 ATGGTTCCTCCATAAATAGGGGCATGAGCCCTTCCTCACAATAA 37398

```
RESULT 15
US-10-172-118-1657
; Sequence 1657, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1657
; LENGTH: 2113
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_016582
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1657

Query Match      93.6%; Score 96.4; DB 6; Length 2113;
Best Local Similarity 99.0%; Pred. No. 3.1e-24;
Matches 97; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  CGGACAGCGGCAGCAGTCCAGCTCTGGTTTCCTTCGTTTATCTCTGTAGATGAA 60
      |  |||||
Db      2016 CGGACAGCGGCAGCAGTCCAGCTCTGGTTTCCTTCGTTTATCTCTGTAGATGAA 60
      |  |||||

Qy      61  ATGTTTCCCATAAATAAGGGGCATGAGCCCTTCCTCAC 98
      |  |||||
Db      2076 ATGTTTCCCATAAATAAGGGGCATGAGCCCTTCCTCCC 2113
```

Search completed: December 6, 2005, 22:57:10
Job time : 257.473 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:34:11 ; Search time 70.31 Seconds
(without alignments)
455.919 Million cell updates/sec

Title: US-09-980-046B-8

Perfect score: 103

Sequence: 1 cggacagcggcagcagtc.....tgagcccttctccacaaaa 103

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New:*

- 1: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/PC1_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
- 9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
- 10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	27.8	27.0	1430	6	US-10-750-185-34214
C 2	26.8	26.0	1976	6	US-10-750-185-24699
C 3	26.6	25.8	1743	6	US-10-750-185-47197
C 4	26.4	25.6	1222	6	US-10-750-185-57246
C 5	26.2	25.4	3883	6	US-10-750-185-61102
C 6	25.8	25.0	1573	6	US-10-750-185-45497
C 7	25.8	25.0	1699	6	US-10-750-185-50782
C 8	25.6	24.9	1737	6	US-10-750-185-47021
C 9	25.6	24.9	1230	6	US-10-750-185-41831
C 10	25.4	24.7	2096	6	US-10-750-185-31080
C 11	25.4	24.7	150450	7	US-11-112-908-54
C 12	25.2	24.5	600	6	US-10-750-185-21418
C 13	25.2	24.5	1943	6	US-10-750-185-31066
C 14	24.8	24.1	1599	6	US-10-750-185-33730
C 15	24.8	24.1	2028	6	US-10-750-185-47588
C 16	24.6	23.9	1187	6	US-10-750-185-32572
C 17	24.6	23.9	1597	6	US-10-750-185-33998
C 18	24.6	23.9	173115	7	US-11-112-908-65
C 19	24.4	23.7	748	6	US-10-750-185-51701
C 20	24.4	23.7	919	6	US-10-750-185-45500
C 21	24.2	23.5	3058	6	US-10-750-185-51651
C 22	24	23.3	963	6	US-10-750-185-25505
C 23	24	23.3	1150	6	US-10-750-185-58765

C 24	23.3	1795	6	US-10-750-185-54055
C 25	23.3	2350	6	US-10-821-234-53
C 26	23.1	966	6	US-10-750-185-38740
C 27	23.1	1331	6	US-10-750-185-61358
C 28	23.1	1560	6	US-10-467-657-3947
C 29	22.9	598	6	US-10-750-185-20207
C 30	22.9	1794	6	US-10-750-185-40377
C 31	22.9	2004	6	US-10-750-185-33449
C 32	22.9	2526	6	US-10-750-185-33449
C 33	22.9	149419	7	US-11-112-908-49
C 34	22.9	161726	7	US-11-112-908-48
C 35	22.9	161726	7	US-11-112-908-52
C 36	22.9	166111	7	US-11-112-908-47
C 37	22.7	450	6	US-10-750-185-55289
C 38	22.7	1169	6	US-10-827-641-115
C 39	22.7	1336	6	US-10-750-185-50334
C 40	22.7	1420	6	US-10-750-185-27685
C 41	22.7	1466	6	US-10-750-185-64081
C 42	22.7	1499	6	US-10-750-185-39776
C 43	22.7	2176	6	US-10-750-185-47006
C 44	22.5	600	6	US-10-750-185-4062
C 45	22.5	1025	6	US-10-750-185-27012

ALIGNMENTS

RESULT 1

US-10-750-185-34214/c
; Sequence 34214, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: WM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34214
; LENGTH: 1430
; TYPE: DNA
; ORGANISM: Bovine 19866881402808
US-10-750-185-34214

Query Match 27.0%; Score 27.8; DB 6; Length 1430;
Best Local Similarity 54.4%; Pred. No. 0.79;
Matches 56; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy	1	CGGACAGCGCAGCAGCTCCGCTCGGTTCTCTTCGTTTATCTGTTCAGATGAA 60
Db	188	CAGACAGACCTCAGCAGCCTCAGCACATAGTAGTGCTCAGGAGGGCTGTCGATGGA 129
Qy	61	ATGTTTCCCAATAATAGGGGCATGAGCCCTTCCTCACAATAA 103
Db	128	TGGGGAACACACAGCAGGTATTATCACTCTCTTCTCCAGAA 86

RESULT 2

US-10-750-185-24699
; Sequence 24699, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.

```
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 24699
; LENGTH: 1976
; TYPE: DNA
; ORGANISM: Bovine 19866880757769
US-10-750-185-24699

Query Match      26.0%; Score 26.8; DB 6; Length 1976;
Best Local Similarity 64.5%; Pred. No. 2.1;
Matches 40; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 36 TCTCGGTTTATCTGTAGAAATGAAATGGTTCCCAATAAATAAGGGCATGAGCCCTTCCT 95
DB 157 TCTCTTGCATTTTGTAAAGATGACAGTCTGACCATAAACAAGAGTATTTTCCTTCCT 216
QY 96 CA 97
DB 217 GA 218

RESULT 3
US-10-750-185-47197
; Sequence 47197, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 47197
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Bovine 19866880234735
US-10-750-185-47197

Query Match      25.8%; Score 26.6; DB 6; Length 1743;
Best Local Similarity 63.1%; Pred. No. 2.3;
Matches 41; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 22 AGCTCTGGTTTCCTTCGCGTTTATCTGTAGAAATGTTCCCAATAAATAAGGGG 81
DB 1668 AGGCTCGGTTTGTGGTGGATTTAAATACATATAATAAATATGTACCTAAATTTGAAG 1727
QY 82 CATGA 86
DB 1728 AATGA 1732

RESULT 4
US-10-750-185-57246
```

```
; Sequence 57246, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 57246
; LENGTH: 1222
; TYPE: DNA
; ORGANISM: Bovine 19866880532291
US-10-750-185-57246

Query Match      25.6%; Score 26.4; DB 6; Length 1222;
Best Local Similarity 59.2%; Pred. No. 2.4;
Matches 45; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 26 CTGGTTTCCTTCCTCGGTTTATCTGTAGAAATGAAATGGTTCCCAATAAATAAGGGCATG 85
DB 716 CTGTTCACTCTGTAGGATTAGTATGAGGTAAACATACACTTGACATAAACAATGCTGAATA 775
QY 86 AGCCCTTCCTCAAAA 101
DB 776 AGCCTTGCTTAATAAA 791

RESULT 5
US-10-750-185-61102
; Sequence 61102, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 61102
; LENGTH: 3883
; TYPE: DNA
; ORGANISM: Bovine 19866881158230
US-10-750-185-61102

Query Match      25.4%; Score 26.2; DB 6; Length 3883;
Best Local Similarity 56.3%; Pred. No. 4.5;
Matches 49; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 16 AGTCCCAGCTCTCGGTTTCCTTCGCGTTTATCTGTAGAAATGAAATGGTTCCCAATAA 75
DB 45 AGTATAAACTCAATTTCTTTTCAAGATACATCTCTTTCCCTGATATCATCTCTTTAACT 104
QY 76 AAGGGGCATGAGCCCTTCCTCAAAAA 102
```

Db 105 CAGGCGCTGGAGGCTTTTAATCAGACAA 131

RESULT 6

US-10-750-185-45497/c
; Sequence 45497, Application US/10750185
; Publication No. US20050260603A1

; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2

; CURRENT APPLICATION NUMBER: US/10/750,185

; CURRENT FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/437,482

; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922

; SOFTWARE: PatentIN version 3.1

; SEQ ID NO 45497

; LENGTH: 1573

; TYPE: DNA

; ORGANISM: Bovine 19866880410534

US-10-750-185-45497

Query Match 25.0%; Score 25.8; DB 6; Length 1573;

Best Local Similarity 53.5%; Pred. No. 4.4;

Matches 54; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 3 GACGACGGCAGCAGCTCCAGCTCGGTTCTCTCGTTTATCTGTTAGAAATGAAT 62

Db 179 GACGAGGGCTTAAATAAGTCCCACTTCTCTGTTGGATTAAATCTGTTGGATTAAATC 120

Qy 63 GGTTCCTCCATAAATAAGGGGCATGAGCCCTTCTCTCAAAAAA 103

Db 119 TGTAAACTAATCTGTTTGATTAAATCATGACCTCCAGAA 79

RESULT 7

US-10-750-185-50782/c

; Sequence 50782, Application US/10750185

; Publication No. US20050260603A1

; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2

; CURRENT APPLICATION NUMBER: US/10/750,185

; CURRENT FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/437,482

; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922

; SOFTWARE: PatentIN version 3.1

; SEQ ID NO 50782

; LENGTH: 1699

; TYPE: DNA

; ORGANISM: Bovine 19866880598678

US-10-750-185-50782

Query Match 25.0%; Score 25.8; DB 6; Length 1699;

Best Local Similarity 63.9%; Pred. No. 4.6;

Matches 39; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 43 TTATTCTGTTAGAAATGTTCCCATAAATAAGGGGCATGAGCCCTTCTCTCAAAAA 102

Db 1632 TTTTCTGTGAGGTTTAAAGCATACCCCAAAATCAGATGACGTGGACCTTCTCTCGAAA 1573

Qy 103 A 103

Db 1572 A 1572

RESULT 8

US-10-750-185-47021

; Sequence 47021, Application US/10750185

; Publication No. US20050260603A1

; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2

; CURRENT APPLICATION NUMBER: US/10/750,185

; CURRENT FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/437,482

; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922

; SOFTWARE: PatentIN version 3.1

; SEQ ID NO 47021

; LENGTH: 757

; TYPE: DNA

; ORGANISM: Bovine 19866880858356

US-10-750-185-47021

Query Match 24.9%; Score 25.6; DB 6; Length 757;

Best Local Similarity 66.1%; Pred. No. 4;

Matches 37; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 11 GCAGCAGTCCAGCTCTGGTTTCTCTCGGTTTATCTGTTAGAAATGAATGGTT 66

Db 562 GCTGAGTCCAGGCTCGGTTTCTCTCATCTGTAAATGGGAGAGATAGATTGTT 617

RESULT 9

US-10-750-185-41831/c

; Sequence 41831, Application US/10750185

; Publication No. US20050260603A1

; GENERAL INFORMATION:

; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2

; CURRENT APPLICATION NUMBER: US/10/750,185

; CURRENT FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/437,482

; PRIOR FILING DATE: 2002-12-31

; NUMBER OF SEQ ID NOS: 64922

; SOFTWARE: PatentIN version 3.1

; SEQ ID NO 41831

; LENGTH: 1230

; TYPE: DNA

; ORGANISM: Bovine 19866881468051

US-10-750-185-41831

Query Match

Best Local Similarity 24.9%; Score 25.6; DB 6; Length 1230;

Matches 46; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 2 GGACAGCGGAGCAGCCAGCTCGTTTCCTTCGCTTATTCGTTAGTAATAA 61
Db 413 GGTGGAAGATGCTTGTCACTCTGTTCTGTTGTTTATTTAGTCTTTAGAGTACT 354
QY 62 TGGTTCCCATAAATAAGGGG 81
Db 353 CCCTTCTCATGAATGAGGTG 334

RESULT 10

US-10-750-185-31080
; Sequence 31080, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 31080
; LENGTH: 2096
; TYPE: DNA
; ORGANISM: Bovine 19866881218394
US-10-750-185-31080

Query Match 24.7%; Score 25.4; DB 6; Length 2096;
Best Local Similarity 58.7%; Pred. No. 6.9;
Matches 44; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 2 GGACAGCGGAGCAGCTCGTTTCCTTCGCTTATTCGTTAGTAATAA 61
Db 137 GTACAGCGGAGAGGCTCAAACTGCTTTCTTTCTAAATTAATAAAAA 196
QY 62 TGGTTCCCATAAATA 76
Db 197 TTTTATTTGAATGA 211

RESULT 11

US-11-112-908-54
; Sequence 54, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 54
; LENGTH: 150450
; TYPE: DNA
; ORGANISM: Homo sapiens

US-11-112-908-54

Query Match 24.7%; Score 25.4; DB 7; Length 150450;
Best Local Similarity 58.7%; Pred. No. 35;
Matches 44; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 27 TGGTTTCCTTCGCTTATTCGTTAGTAATAAATGGTTCCCATAAATAAGGGGCATGA 86
Db 37550 TGTGATGCTTCAGCTTGTCTTTTGGCTTAAATTCCTCCACCTATGAGGGTCTTTT 37609
QY 87 GCCCTTCCTCACAA 101
Db 37610 GTGGTTCCATATAA 37624

RESULT 12

US-10-750-185-21418/c
; Sequence 21418, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 21418
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT10387
US-10-750-185-21418

Query Match 24.5%; Score 25.2; DB 6; Length 600;
Best Local Similarity 56.2%; Pred. No. 5.1;
Matches 45; Conservative 1; Mismatches 34; Indels 0; Gaps 0;
QY 2 GGACAGCGGAGCAGCTCCAGCTCGTTTCCTTCGCTTATTCGTTAGTAATAA 61
Db 113 GGTGGAAGGATGCTTGTCACTCTGTTCTGTTTATTTAGTCTTTAGAGTACT 54
QY 62 TGGTTCCCATAAATAAGGGG 81
Db 53 CCCTTCTCATGAATGAGTG 34

RESULT 13

US-10-750-185-31066/c
; Sequence 31066, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922

```
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 31066
; LENGTH: 1943
; TYPE: DNA
; ORGANISM: Bovine 19866880870888
US-10-750-185-31066

Query Match      24.5%; Score 25.2; DB 6; Length 1943;
Best Local Similarity 57.7%; Pred. No. 8;
Matches 45; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy  2  GGACAGAGCGGAGAGTCCAGCTCTGTTTCCTTCGCGTTTATTCGTTAGAAATGAAA 61
Db  1128 GGACAGAGTCCAGAGTGGGGGACCGCATACCAATTGTTGGAAGCGTCTCTTAGGATGAAA 1069

Qy  62  TGGTTCACATAAATAAGG 79
Db  1068 TTGTGTCGTAGGCAGGG 1051
```

```
RESULT 14
US-10-750-185-33730/c
; Sequence 33730, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 33730
; LENGTH: 1599
; TYPE: DNA
; ORGANISM: Bovine 19866880783812
US-10-750-185-33730
```

```
Query Match      24.1%; Score 24.8; DB 6; Length 1599;
Best Local Similarity 56.0%; Pred. No. 10;
Matches 47; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy  13  AGCAGTCCAGCTCTGGTTTCTTCGTTTATTCGTTAGAAATGTTCCCATTA 72
Db  949  ATCACTCCAGCTCTACAAAATTCATGGCTCCCTCTTCACTGTGAAAGGCTTGCCCAA 890

Qy  73  AATAGGGGGCATGAGCCCTTCCTC 96
Db  889  AGTGAGAGCCTCAGCTCTGCCCC 866
```

```
RESULT 15
US-10-750-185-47588/c
; Sequence 47588, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
```

```
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 47588
; LENGTH: 2028
; TYPE: DNA
; ORGANISM: Bovine 19866881295675
US-10-750-185-47588

Query Match      24.1%; Score 24.8; DB 6; Length 2028;
Best Local Similarity 80.6%; Pred. No. 11;
Matches 29; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy  29  GTTTCCTTCTCGGTTTATTCGTTAGAAATGAAATGG 64
Db  1892 GTTTCATCTGTTATTTGTTCTGTTAAAAATGAAATGG 1857

Search completed: December 6, 2005, 23:16:54
Job time : 70.31 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 47.1624 Seconds
(without alignments)
4239.411 Million cell updates/sec

Title: US-09-980-046B-9

Perfect score: 30

Sequence: 1 cggaaataaaggctgtgtaagacaaaaa 30

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq_21:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

14: Geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	30	4	AAC89740
2	28.4	94.7	547	4	Aaf75084 Human gas
3	28.4	94.7	642	8	Abx63536 Human CDN
4	28.4	94.7	668	2	Aaz77534 Human ova
5	28.4	94.7	670	2	Aaz33612 Human bre
6	28.4	94.7	678	12	Adq17878 Human sof
7	28.4	94.7	1323	6	Abg54427 Human ova
8	28.4	94.7	1389	12	Adq22502 Human sof
9	26.4	88.0	308	13	Adt50717 Cancer re
10	24.4	81.3	509	13	Adul3731 Solid tum
c 11	24	80.0	469	12	Ados5031 Gene #128
12	24	80.0	1036	12	Adi57714 Human bre
c 13	23	76.7	301	3	Aac66000 Human lun
c 14	23	76.7	301	6	Abi49219 Human lun
c 15	23	76.7	301	6	Abq92405 Human lun
c 16	23	76.7	301	9	Ada28394 Human lun
c 17	23	76.7	301	10	Adh36958 Human lun
c 18	23	76.7	301	12	Adm56761 Human lun
c 19	23	76.7	301	14	Adu98453 Lung tumo

c	20	23	76.7	301	14	AEB10261	Aeb10261 Cancer re
c	21	21.2	70.7	535	4	AAL12750	Aal12750 Human bre
c	22	21.2	70.7	905	11	ACN82890	Acn82890 Breast ca
23	21	21	70.0	464	4	AAK88758	Aak88758 Human dig
24	21	21	70.0	464	5	AAS31792	Aas31792 Human liv
25	21	21	70.0	464	6	ABN90147	Abn90147 Human liv
26	21	21	70.0	464	11	ADJ14914	Adj14914 Human liv
27	21	21	70.0	486	4	AAK91075	Aak91075 Human dig
28	21	21	70.0	486	4	AAK91076	Aak91076 Human dig
29	21	21	70.0	486	5	AAS32111	Aas32111 Human liv
30	21	21	70.0	486	5	AAS32110	Aas32110 Human liv
31	21	21	70.0	486	6	ABN90465	Abn90465 Human liv
32	21	21	70.0	486	6	ABN90466	Abn90466 Human liv
33	21	21	70.0	486	11	ADJ15379	Adj15379 Human liv
34	21	21	70.0	486	11	ADJ15378	Adj15378 Human liv
35	21	21	70.0	634	6	ABL63301	Abi63301 Breast ca
36	21	21	70.0	634	6	ABL62874	Abi62874 Breast ca
37	21	21	70.0	634	6	ABL63092	Abi63092 Breast ca
38	21	21	70.0	634	6	ABK35563	Abk35563 Gene ISG1
39	21	21	70.0	634	6	ABT10900	Abt10900 Human bre
40	21	21	70.0	634	6	ABZ35078	Abz35078 Human gen
41	21	21	70.0	634	6	ABK83926	Abk83926 Human CDN
42	21	21	70.0	634	8	ACC51007	Acc51007 Human bla
43	21	21	70.0	634	8	ACC50195	Acc50195 Breast ca
44	21	21	70.0	634	8	ACC51210	Acc51210 Human Plk
45	21	21	70.0	634	10	ADD19019	Add19019 Human dis

ALIGNMENTS

RESULT 1

AAC89740

ID AAC89740 standard; cDNA; 30 BP.

AC AAC89740;

XX 12-MAR-2001 (first entry)

XX Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 9.

XX Human; cytostatic; immunomodulator; immunostimulant; vulnery;

XX anti-inflammatory; neuroprotective; antibacterial; gene therapy;

XX gastrointestinal inflammation; immune system disorder; genetic disorder;

XX cancer; autoimmune disorder; infection; wound healing; ss.

XX Homo sapiens.

XX WO2000073324-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015191.

XX 01-JUN-1999; 99US-0137058P.

XX (DIGI-) DIGITAL GENE TECHNOLOGIES INC.

XX Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Haseel KW;

XX WPI; 2001-061508/07.

XX New polynucleotides and polypeptides, useful in gene therapy and in

XX diagnosing a pathological condition, e.g. for modulating gene expression

XX disorders.

XX Claim 1; Page 85; 108pp; English.

XX The present sequence is one of a number of isolated human polynucleotides

XX which are useful in gene therapy, and for diagnosing a pathological

XX condition or a susceptibility to it. In particular, the polynucleotides

XX are useful for modulating gene expression in gastrointestinal

CC inflammation. The polynucleotides are useful for chromosome
 CC identification, controlling gene expression through triple helix
 CC formation or antisense DNA or RNA, or identifying individuals from minute
 CC biological samples using DNA-based identification techniques. The
 CC polynucleotides can also be used as an alternative to restriction
 CC fragment length polymorphism (RFLP), by determining the actual base-by-
 CC base DNA sequences of selected portions of an individual's genome. The
 CC polynucleotides may also be used as molecular weight markers on Southern
 CC gels, as diagnostic probes for the presence of a specific mRNA, as a
 CC probe to subtract-out known sequences in the process of discovering novel
 CC polynucleotides, or as an antigen to elicit an immune response. The
 CC polypeptides are useful in diagnostic procedures to detect a disorder.
 CC The polynucleotides and polypeptides are useful for preventing, treating
 CC or ameliorating immune system disorders, genetic disorders, cancers, some
 CC autoimmune disorders, or infections. The polynucleotides and polypeptides
 CC are also useful for differentiating, proliferating or attracting cells,
 CC leading to the regeneration of tissues, especially in wounds or burns.
 CC The polypeptides and polynucleotides may also be used as a food additive
 CC or preservative
 CC
 CC Sequence 30 BP; 15 A; 3 C; 7 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 4; Length 30;

Best Local Similarity 100.0%; Pred. No. 0.033;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGACAAAA 30
 |||||
 Db 1 CGGAATAAAGGCTGTTGTAAGACAAAA 30

RESULT 2

AAAF75084
 ID AAF75084 standard; cDNA; 547 BP.

XX AAF75084;

DT 10-MAY-2001 (first entry)

DE Human colon associated protein cDNA sequence #8.

XX Human; colon; cancer; disease; sa.

XX Homo sapiens.

XX WO200112781-A1.

XX 22-FEB-2001.

XX 11-AUG-2000; 2000WO-US022157.

XX 13-AUG-1999; 99US-0148680P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2001-147551/15.

XX Nucleic acids encoding 13 human colon cancer associated polypeptides,
 PT useful for preventing, diagnosing and/or treating e.g. cancers
 PT (especially colon cancer), Parkinson's disease and diabetic retinopathy.

XX Claim 1; Page 310-311; 326pp; English.

XX The present invention relates to 13 human colon cancer-associated
 CC proteins. These proteins and the nucleic acid encoding them may be used
 CC in the prevention, diagnosis and treatment of diseases associated with
 CC inappropriate colon cancer-associated protein expression

XX Sequence 547 BP; 139 A; 152 C; 180 G; 76 T; 0 U; 0 Other;

XX Query Match 94.7%; Score 28.4; DB 4; Length 547;

Best Local Similarity 96.7%; Pred. No. 0.19;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGACAAAA 30
 |||||
 Db 488 CGGAATAAAGGCTGTTGTAAGACAAAA 517

RESULT 3

ABX63536

ID ABX63536 standard; cDNA; 642 BP.

XX AC ABX63536;

XX 26-FEB-2003 (first entry)

XX Human cDNA #536 differentially expressed in activated vascular tissue.

XX Human; gene; sa; vascular tissue; cytostatic; atherosclerosis; cardiast;

XX hypotensive; antidiabetic; gynaecological; vasotropic; cerebroprotective;

XX gene therapy; vascular disease; cancer; coronary; artery disease;

XX hypertension; diabetes; pre-eclampsia; restenosis;

XX ischaemia-reperfusion injury; stroke.

XX Homo sapiens.

XX US2002137081-A1.

XX 26-SEP-2002.

XX 08-JAN-2002; 2002US-00044090.

XX 28-JUL-2000; 2000US-0222469P.

XX 08-JAN-2001; 2001US-0260483P.

XX (BAND/) BANDMAN O.

XX Bandman O;

XX WPI; 2003-110597/10.

XX Combination for diagnosing, staging, treating, or monitoring the
 PT progression of treatment of a vascular disease, e.g. atherosclerosis,
 PT comprises several cDNAs that are differentially expressed in activated
 PT vascular tissue.

XX Claim 1; Page; 18pp; English.

XX This invention relates to a combination comprising several cDNAs that are
 CC differentially expressed in activated vascular tissue. The invention also
 CC discloses a high throughput method for detecting differentially expressed
 CC cDNAs in a sample. The cDNAs of the invention may have
 CC antiatherosclerotic; cytostatic; cardiast; hypotensive; antidiabetic;
 CC gynaecological; vasotropic and cerebroprotective activities and may be
 CC used in gene therapy. The cDNAs of the invention may be used in a high-
 CC throughput methods for detecting differential expression of one or more
 CC cDNAs in a sample, or screening several molecules or compounds to
 CC identify a molecule or compound that specifically binds a cDNA of the
 CC invention. A protein encoded by the cDNA may be used to screen several
 CC molecules or compounds to identify a ligand that specifically binds to
 CC the protein, or to produce or purify an antibody to the protein that can
 CC be used to detect a protein in a sample or purify a natural or
 CC recombinant protein from a sample. The nucleotides may be useful for
 CC diagnosing, staging, treating, or monitoring the progression of treatment
 CC of a vascular disease, e.g. atherosclerosis, cancer, coronary artery
 CC disease, hypertension, diabetes, pre-eclampsia, ischaemia-reperfusion
 CC injury, restenosis, or stroke. The cDNAs can also be used for large-scale
 CC genetic or gene expression analysis of several new nucleic acid
 CC molecules. Antibodies to the proteins encoded by the cDNAs are useful for
 CC diagnosing pre-pathologic disorders, and chronic or acute diseases
 CC associated with abnormalities in the expression, amount or distribution
 CC of the protein. The present sequence represents a cDNA of the invention
 CC that is differentially expressed in activated vascular tissue. Note: The

CC sequence data for this patent did not form part of the specification, but
CC was obtained in electronic format directly from USPTO at
CC <http://seqdata.uspto.gov/sequence.html?DocID=20020137081>

XX SQ Sequence 642 BP; 141 A; 188 C; 219 G; 94 T; 0 U; 0 Other;

Query Match 94.7%; Score 28.4; DB 8; Length 642;
Best Local Similarity 96.7%; Pred. No. 0.19;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTTAAAGACAAAA 30
Db 606 CGGAATAAAGGCTGTTGTTAAAGACAAAA 635

RESULT 4

AAZ77534
ID AAZ77534 standard; cDNA; 668 BP.

XX AC AAZ77534;

XX DT 10-APR-2000 (first entry)

XX DE Human ovarian tumor cDNA library derived EST fragment 85.

XX KW Expressed sequence tag; EST; human; ovarian tumor; anticancer;

XX KW gene therapy; treatment; ss.

XX OS Homo sapiens.

XX PN DE19817557-A1.

XX PD 21-OCT-1999.

XX PF 09-APR-1998; 98DE-01017557.

XX PR 09-APR-1998; 98DE-01017557.

XX PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;

XX DR WPI; 1999-591920/51.

XX PT New nucleic acid sequences expressed in ovarian, and some other, cancer
PT tissues, and derived polypeptides, for treatment of ovarian cancer and
PT identification of therapeutic agents.

XX PS Claim 3; Page 212; 310pp; German.

XX CC This invention describes novel nucleic acid (cDNA) sequences (A) which
CC have anticancer activity and are highly expressed in ovarian tumor tissue
CC (and some also in testis and breast cancer tissue). The products of the
CC invention can be used for gene therapy. (A) are used (i) for recombinant
CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)
CC are used (i) to identify agents suitable for treatment of ovarian cancer;
CC (ii) directly for treating this form of cancer (including expression from
CC gene therapy vectors) and (iii) for generation of specific antibodies.
CC (A) are identified by assembling ESTs (expressed sequence tags) from a
CC particular tissue type before comparison of expression patterns. This
CC allows a significantly longer fragment of the gene to be revealed, so
CC should reduce the number of failures associated with the fact that ESTs
CC from different libraries may represent different parts of the same
CC unknown gene, distorting the estimated frequency of occurrence in a
CC particular tissue. AAZ77450-277572 represent the human ovarian tumor cDNA
CC library derived EST fragments described in the method of the invention
CC and encode the protein fragments represented in AAY76505-Y76638

XX SQ Sequence 668 BP; 140 A; 197 C; 233 G; 98 T; 0 U; 0 Other;

Query Match 94.7%; Score 28.4; DB 2; Length 668;
Best Local Similarity 96.7%; Pred. No. 0.19;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTTAAAGACAAAA 30
Db 637 CGGAATAAAGGCTGTTGTTAAAGACAAAA 666

RESULT 5

AAZ33612
ID AAZ33612 standard; cDNA; 670 BP.

XX AC AAZ33612;

XX DT 08-DEC-1999 (first entry)

XX DE Human breast tumour-associated EST 2.

XX KW Expressed sequence tag; EST; human; breast; cancer; gene therapy;

XX KW treatment; tumour; cytostatic; medicament; ss.

XX OS Homo sapiens.

XX PN DE19813839-A1.

XX PD 23-SEP-1999.

XX PF 20-MAR-1998; 98DE-01013839.

XX PR 20-MAR-1998; 98DE-01013839.

XX PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;

XX DR WPI; 1999-528981/45.

XX PT Human nucleic acid sequences and protein products from tumor breast
PT tissue, useful for breast cancer therapy.

XX PS Claim 3; 83; 188pp; German.

XX CC This invention describes novel human nucleic acid sequences from tumor
CC breast tissue which have cytostatic activity. The nucleic acid sequences
CC can be used to produce and isolate full-length gene sequences. They can
CC be used to express proteins, which can be used as tools to find an
CC activity against breast cancer. The sequences can be used in sense or
CC antisense form. They are especially useful for medicaments for gene
CC therapy to treat breast cancer. AAZ33611-248617 represents expressed
CC sequence tags described in the method of the invention

XX SQ Sequence 670 BP; 142 A; 197 C; 233 G; 98 T; 0 U; 0 Other;

Query Match 94.7%; Score 28.4; DB 2; Length 670;
Best Local Similarity 96.7%; Pred. No. 0.19;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTTAAAGACAAAA 30
Db 637 CGGAATAAAGGCTGTTGTTAAAGACAAAA 666

RESULT 6

ADQ17878
ID ADQ17878 standard; DNA; 678 BP.

XX AC ADQ17878;

XX DT 26-AUG-2004 (first entry)

XX DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 695.

XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
XX ds.

```

OS Homo sapiens.
XX
XX PN W02004048938-A2.
XX
XX PD 10-JUN-2004.
XX
XX PF 26-NOV-2003; 2003WO-US038193.
XX
XX PR 26-NOV-2002; 2002US-0429739P.
XX
XX PA (PROT-) PROTEIN DESIGN LABS INC.
XX
XX PI Aziz N, Gineburg WM, Zlotnik A;
XX
XX DR WPI; 2004-441208/41.
XX
XX PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
XX PS Example 2; SEQ ID NO 695; 210pp; English.
XX
XX CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytosstatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
XX SQ Sequence 678 BP; 158 A; 189 C; 227 G; 104 T; 0 U; 0 Other;

Query Match 94.7%; Score 28.4; DB 12; Length 678;
Best Local Similarity 96.7%; Pred. No. 0.19;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTTGTAAGACACAAAA 30
Db |||||
620 CGGAATAAAGGCTGTTGTAAGACACAAAA 649

RESULT 7
ABQ54427
ID ABQ54427 standard; cDNA; 1323 BP.
XX
XX AC ABQ54427;
XX
XX DT 22-AUG-2002 (first entry)
XX
XX DE Human ovarian antigen HPAMB04 cDNA, SEQ ID NO:307.
XX
XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
XX inflammatory condition; immune disorder; blood disorder;
XX cardiovascular disorder; respiratory disorder; neurological disorder;
XX gastrointestinal disorder; urinary system disorder; drug screening;
XX gene therapy; chromosome mapping; forensic analysis;
XX antibody preparation; cytostatic; immunomodulatory; neuroprotective;
XX antiinflammatory; gynaecological; reproductive; gene; ss.
XX
XX OS Homo sapiens.
XX
XX PN W0200200677-A1.
XX
XX XX

```

```

PD 03-JAN-2002.
XX
XX PF 07-JUN-2001; 2001WO-US018569.
XX
XX PR 07-JUN-2000; 2000US-0209467P.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Birse CE, Rosen CA;
XX
XX DR WPI; 2002-147878/19.
XX
XX DR P-PSDB; ABP41350.
XX
XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
XX cancer), immune disorders, cardiovascular disorders and neurological
XX diseases.
XX
XX PS Claim 1; SEQ ID NO 307; 2922pp; English.
XX
XX CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to
XX recombinant vectors and host cells comprising human ovarian antigen
XX polynucleotides, antibodies against human ovarian antigens, and the use
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX treating, prognosing or preventing various ovary and/or breast-related
XX disorders. Such conditions include ovarian cancer and breast cancer, and
XX metastatic tumours of ovarian or breast origin, reproductive system
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX vaginitis), immune disorders (e.g., congenital and acquired
XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX respiratory disorders, neurological disorders, gastrointestinal disorders
XX and urinary system disorders. Ovarian antigen polypeptides and
XX polynucleotides may also be used in screening for compounds which
XX modulate ovarian antigen expression or activity. The polynucleotides may
XX further be used for gene therapy, chromosome mapping, in the
XX identification of individuals and in forensic analysis, and the
XX polypeptides may be used as food additives or to prepare antibodies
XX useful in disease diagnosis, drug targeting and phenotyping. The present
XX sequence represents cDNA encoding a human ovarian antigen of the
XX invention. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 1323 BP; 343 A; 349 C; 394 G; 235 T; 0 U; 2 Other;

Query Match 94.7%; Score 28.4; DB 6; Length 1323;
Best Local Similarity 96.7%; Pred. No. 0.2;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTTGTAAGACACAAAA 30
Db |||||
659 CGGAATAAAGGCTGTTGTAAGACACAAAA 688

RESULT 8
ADQ22502
ID ADQ22502 standard; DNA; 1389 BP.
XX
XX AC ADQ22502;
XX
XX DT 26-AUG-2004 (first entry)
XX
XX DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 5322.
XX
XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
XX ds.

```

```
XX OS Homo sapiens.
XX FN WO2004048938-A2.
XX PD 10-JUN-2004.
XX PF 26-NOV-2003; 2003WO-US038193.
XX PR 26-NOV-2002; 2002US-0429739P.
XX PA (PROT-) PROTEIN DESIGN LABS INC.
XX PI Aziz N, Ginsburg WM, Zlotnik A;
XX DR WPI; 2004-441208/41.
XX PT Early detection of soft tissue sarcoma comprises determining expression
XX of a gene in a first soft tissue sample and a normal soft tissue sample
XX and comparing the gene expression, also useful in treating soft tissue
XX sarcoma.
XX PS Example 2; SEQ ID NO 5322; 210pp; English.
XX CC The invention relates to a novel method for detecting soft tissue sarcoma
XX which comprises obtaining a first soft tissue sample from an individual
XX and a normal soft tissue sample from the same or different individual,
XX determining the expression of a gene in both samples and comparing the
XX expression of the gene in both soft tissue samples, where a higher level
XX of protein expression in the first soft tissue sample indicates the
XX presence of soft tissue sarcoma. The method of the invention has
XX cytosstatic applications and may be useful for detecting soft tissue
XX sarcoma, possibly via gene therapy or vaccine production. The nucleic
XX acid sequences may be useful in diagnostic and screening applications.
XX The current sequence is that of a human soft tissue sarcoma-upregulated
XX DNA of the invention. The current sequence is not shown within the
XX specification per se but was submitted in CD format by the inventor.
XX SQ Sequence 1389 BP; 483 A; 300 C; 367 G; 239 T; 0 U; 0 Other;

Query Match 94.7%; Score 28.4; DB 12; Length 1389;
Best Local Similarity 96.7%; Pred. No. 0.2;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTGTAAGACAAAAA 30
Db 554 CGGAATAAAGGCTGTGTAAGACAAAAA 583

RESULT 9
ADT50717
ID ADT50717 standard; DNA; 308 BP.
AC ADT50717;
DT 13-JAN-2005 (first entry)
XX Cancer related nucleic acid sequence #21.
XX ds; gene; cytostatic; gene therapy; vaccine; diagnosis; breast; colon;
XX lung; ovarian; prostate; cancer.
XX OS Homo sapiens.
XX FN WO2004092338-A2.
XX PD 28-OCT-2004.
XX PF 12-APR-2004; 2004WO-US011104.
XX PR 11-APR-2003; 2003US-0462399P.
XX PR 01-JUL-2003; 2003US-0484339P.
XX CC

PA (DIAD-) DIADEXUS INC.
XX Macina RA, Turner LR, Sun Y, Tam A;
XX WPI; 2004-766851/75.
XX New cancer specific nucleic acid (CaSNA) molecules, useful for
XX diagnosing, monitoring the presence of, or treating a patient with
XX breast, colon, lung, ovarian, or prostate cancer.
XX Claim 1; SEQ ID NO 21; 891pp; English.
XX The invention relates to an isolated nucleic acid molecule (I)
XX selectively hybridizing to, or comprising at least 95% sequence identity
XX to, any of the 362 nucleotide sequences fully defined in the
XX specification. The nucleic acid molecules and polypeptides are useful for
XX diagnosing, monitoring the presence of, or treating a patient with
XX breast, colon, lung, ovarian, or prostate cancer. This sequence
XX corresponds to a nucleic acid of the invention.
XX SQ Sequence 308 BP; 68 A; 87 C; 110 G; 43 T; 0 U; 0 Other;

Query Match 88.0%; Score 26.4; DB 13; Length 308;
Best Local Similarity 96.4%; Pred. No. 1.1;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTGTAAGACAAA 28
Db 281 CGGAATAAAGGCTGTGTAAGAGAAA 308

RESULT 10
ADU13731
ID ADU13731 standard; DNA; 509 BP.
AC ADU13731;
DT 27-JAN-2005 (first entry)
XX Solid tumour prognosis gene seqid 4170.
XX cytostatic; gene therapy; expression profile; solid tumour;
XX peripheral blood mononuclear cell; PBMC; prognosis; ds.
XX OS Unidentified.
XX FN WO2004097052-A2.
XX PD 11-NOV-2004.
XX PF 29-APR-2004; 2004WO-US013587.
XX PR 29-APR-2003; 2003US-0466067P.
XX PR 23-JAN-2004; 2004US-0538246P.
XX PA (AMHP ) WYETH.
XX PA (STRA/) STRAHS A.
XX Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
XX Immerman F, Dörner AJ;
XX WPI; 2004-804779/79.
XX A method, useful for prognosing and treating solid tumor, comprises
XX comparing an expression profile of a gene expressed in peripheral blood
XX mononuclear cells to a reference expression profile of a gene.
XX Disclosure; Page; 111pp; English.
XX The invention describes a method comprising comparing an expression
XX profile of at least one gene in a peripheral blood sample of a patient to
XX at least one reference expression profile of the at least one gene, where
XX the patient has a solid tumour, and each of the gene is differentially
```

CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
 CC of patients as compared to PBMCs of a second class of patients, where
 CC both of the first and second classes of patients have the solid tumour, and
 CC each of the first and second classes is a subcluster formed by an
 CC unsupervised clustering analysis of gene expression profiles in PBMCs of
 CC a population of patients who have the solid tumour, and where the
 CC majority of the first class of patients has a first clinical outcome, and
 CC the majority of the second class of patients has a second clinical
 CC outcome. Also described are: a system comprising (i) a memory or a
 CC storage medium including data that represent an expression profile of at
 CC least one gene in a peripheral blood sample of a patient who has a solid
 CC tumour, (ii) at least another storage medium including data that
 CC represent at least one reference expression profile of the gene, (iii) a
 CC program capable of comparing the expression profile to the reference
 CC expression profile, and (iv) a processor capable of executing the
 CC program, where expression levels of the gene in peripheral blood
 CC mononuclear cells of patients who have the solid tumour correlate with
 CC clinical outcomes of the patients; and a nucleic acid or protein array
 CC comprising concentrated probes for solid tumour prognosis genes, where
 CC each of the solid tumour prognosis genes is differentially expressed in
 CC PBMCs of a first class of patients as compared to PBMCs of a second class
 CC of patients, where both the first and second classes of patients have a
 CC solid tumour, and where the first class of patients has a first clinical
 CC outcome, and the second class of patients has a second clinical outcome.
 CC The method, system, and array are useful for prognosing and treating
 CC solid tumours. This sequence represents a solid tumour prognosis gene of
 CC the invention. Note: the sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 509 BP; 120 A; 95 C; 128 G; 74 T; 0 U; 92 Other;

Query Match 81.3%; Score 24.4; DB 13; Length 509;
 Best Local Similarity 83.3%; Pred. No. 6.9;
 Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGACAAAA 30
 |||||
 Db 236 CGGAATAAAGGCTGTTGTAANNNGAAAA 265

RESULT 11

AD055031/c

ID AD055031 standard; DNA; 469 BP.

XX AC AD055031;

XX DT 15-JUL-2004 (first entry)

XX DE Gene #128 with increased gene expression in renal cell carcinoma.

XX KW ds; gene; cytostatic; gene therapy; differential expression;

XX KW renal cell carcinoma; clear cell RCC; papillary RCC;

XX KW chromophobe/oncocytoma RCC; sarcomatoid RCC; TCC; Wilms' tumor;

XX KW gene expression; kidney cancer; diagnostic marker; cancer.

XX OS Homo sapiens.

XX PN WO2004032842-A2.

XX PD 22-APR-2004.

XX PF 06-OCT-2003; 2003WO-US031476.

XX PR 04-OCT-2002; 2002US-0415775P.

XX PA (VAND-) VAN ANDEL INST.

XX PI Teh BT, Takahashi M;

XX XX WPI; 2004-340789/31.

XX PT New nucleic acid and polypeptide compositions, useful in the field of

PT molecular biology and medicine, in particular for gene expression
 PT profiling, identifying diagnostic markers, and treating certain types of
 PT kidney cancer.

XX Example IV; SEQ ID NO 128; 53pp; English.

XX The invention relates to novel genes that are differentially expressed in
 CC subtypes of renal cell carcinomas and methods of detecting them using
 CC nucleic acids and probes. The nucleic acid probes hybridize with part or
 CC all of a coding sequence that is overexpressed in clear cell renal cell
 CC carcinoma (CC-RCC), papillary RCC, chromophobe/oncocytoma RCC,
 CC sarcomatoid RCC, TCC, or Wilms' tumors, which overexpression is based on
 CC comparison to a baseline value. The methods and compositions of the
 CC present invention are useful in the field of molecular biology and
 CC medicine, in particular for gene expression profiling of certain types of
 CC kidney cancer, in identifying diagnostic markers, and treating such
 CC cancer patients. This sequence corresponds to a gene with increased
 CC expression in CC-RCC.

SQ Sequence 469 BP; 69 A; 163 C; 139 G; 98 T; 0 U; 0 Other;

Query Match 80.0%; Score 24; DB 12; Length 469;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24

Db 25 CGGAATAAAGGCTGTTGTAAGA 2

RESULT 12

AD157714

ID AD157714 standard; cDNA; 1036 BP.

XX AC AD157714;

XX DT 22-APR-2004 (first entry)

XX DE Human breast specific nucleic acid (BSNA) #85.

XX KW Human; breast specific nucleic acid; BSNA; gene; ss; metastasis;

XX KW breast cancer; cytostatic.

XX OS Homo sapiens.

XX PN WO2003106648-A2.

XX PD 24-DEC-2003.

XX PF 16-JUN-2003; 2003WO-US018934.

XX PR 14-JUN-2002; 2002US-0389327P.

XX PA (DIAD-) DIADEXUS INC.

XX PI Salceda S, Macina RA, Turner LR, Sun Y, Liu C;

XX DR WPI; 2004-082185/08.

XX Novel isolated polypeptide comprising breast specific protein sequences,
 PT useful for diagnosing or monitoring presence and metastases of breast
 PT cancer in patient.

XX Claim 1; SEQ ID NO 85; 370pp; English.

XX The invention relates to human breast specific nucleic acids (BSNA) and
 CC the breast specific proteins (BSP) they encode. The nucleic acids are
 CC useful for determining the presence of a BSNA in a sample which involves
 CC contacting the sample with a BSNA under conditions in which the BSNA will
 CC selectively hybridise to a BSNA in the sample, and detecting the
 CC hybridisation. The nucleic acids are useful for determining the presence
 CC of a BSP in a sample which involves contacting the sample with suitable
 CC reagent under conditions in which the reagent will selectively interact

CC with the BSP, and detecting the interaction of the reagent with a BSP in
CC the sample. The nucleic acids and proteins are useful for diagnosing or
CC monitoring the presence and metastases of breast cancer in a patient,
CC which involves determining an amount of nucleic acid or protein and
CC comparing the determined amount of nucleic acid or protein in the sample
CC of the patient to the amount of a breast specific marker in a normal
CC control, where a difference in the determined amount in the sample
CC compared to the amount in the control is associated with the presence of
CC breast cancer. The sequences are useful for treating a patient with
CC breast cancer, involving administering a composition consisting of a BSNA
CC or a BSP to a patient, where the administration induces an immune
CC response against the breast cancer cell expressing the BSNA or BSP. This
CC sequence represents a human BSNA of the invention.

XX SQ Sequence 1036 BP; 199 A; 298 C; 369 G; 170 T; 0 U; 0 Other;

Query Match 80.0%; Score 24; DB 12; Length 1036;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTGTAAGA 24
|||||
DB 1012 CGGAATAAAGGCTGTGTAAGA 1035

RESULT 13
AAC66000/c
ID AAC66000 standard; cDNA; 301 BP.

XX AAC66000;

XX 21-FEB-2001 (first entry)

XX Human lung cancer-associated cDNA clone 25405.

XX Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
XX vaccine; detection; ss.

XX Homo sapiens.

XX WO200061612-A2.

XX 19-OCT-2000.

XX 03-APR-2000; 2000WO-US008896.

XX 02-APR-1999; 99US-00285479.

XX 17-DEC-1999; 99US-00466396.

XX 30-DEC-1999; 99US-00476496.

XX 10-JAN-2000; 2000US-00480884.

XX 22-FEB-2000; 2000US-00510376.

XX (CORI-) CORIXA CORP.

XX Wang T, Fan L;

XX WPI; 2000-628399/60.

XX Isolated polypeptide comprising an immunogenic portion of a lung tumor
XX protein is used for detecting and monitoring progression of lung cancer
XX in a patient.

XX Claim 25a; Page 229; 261pp; English.

XX This invention describes a novel isolated polypeptide (I) which
XX comprising an immunogenic portion of a lung tumor protein or variant (P2)
XX which have cytostatic activity. The polypeptides and polynucleotides are
XX used in compositions and vaccines to inhibit the development of cancer,
XX especially lung cancer, in a patient. Methods described in the invention
XX can be used to monitor the progression of a cancer by carrying out the
XX detection at subsequent time points and comparing the results from the
XX different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
XX are treated with P2, polynucleotides encoding P2 or antigen presenting

CC cells expressing P2 and then administered to the patient to inhibit
CC development of cancer

XX SQ Sequence 301 BP; 41 A; 109 C; 81 G; 67 T; 0 U; 3 Other;

Query Match 76.7%; Score 23; DB 3; Length 301;
Best Local Similarity 95.8%; Pred. No. 23;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTGTAAGA 24
|||||
DB 25 CGGAATAAAGGCTGTGTAAGA 2

RESULT 14

ABL49219/c

ID ABL49219 standard; cDNA; 301 BP.

XX ABL49219;

XX 01-MAY-2002 (first entry)

XX Human lung tumour cDNA sequence clone 25405 SEQ ID NO:304.

XX Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
XX immune response; ss.

XX Homo sapiens.

XX WO200200174-A2.

XX 03-JAN-2002.

XX 28-JUN-2001; 2001WO-US021065.

XX 28-JUN-2000; 2000US-00606421.

XX 02-AUG-2000; 2000US-00630940.

XX 21-AUG-2000; 2000US-00643597.

XX 15-SEP-2000; 2000US-00662786.

XX 09-OCT-2000; 2000US-00685696.

XX 12-DEC-2000; 2000US-00735705.

XX 07-MAY-2001; 2001US-00850716.

XX (CORI-) CORIXA CORP.

XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
XX McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
XX Vedvick TS, Carter D, Watanabe Y, Peckham DW;

XX WPI; 2002-090513/12.

XX Polynucleotides encoding lung tumor polypeptides, useful for treating
XX lung cancer or stimulating an immune response.

XX Example 1; Page 306; 374pp; English.

XX The present invention describes human lung tumour proteins. Human lung
XX tumour proteins and polynucleotides have cytostatic and immunostimulant
XX activities, and can be used in vaccine production. Compositions
XX comprising the lung tumour proteins, polynucleotides, antibodies, fusion
XX proteins, T cell populations, or antigen presenting cells that express
XX the lung tumour proteins are useful for treating lung cancer or
XX stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
XX ABB75070 represent sequences used in the exemplification of the present
XX invention

XX SQ Sequence 301 BP; 41 A; 109 C; 81 G; 67 T; 0 U; 3 Other;

Query Match 76.7%; Score 23; DB 6; Length 301;
Best Local Similarity 95.8%; Pred. No. 23;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGGCTGTGTAAGA 24

```

Db      25  CGNAATAAAGGCTGTTGTAAGA 2
|||||
RESULT 15
ABQ92405/c
ID      ABQ92405 standard; cDNA; 301 BP.
XX
XX      ABQ92405;
XX
XX      07-OCT-2002 (first entry)
XX
XX      Human lung cancer associated cDNA sequence SEQ ID NO:304.
XX
XX      Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine; gene;
XX      ss.
XX
XX      Homo sapiens.
XX
XX      WO200247534-A2.
XX
XX      20-JUN-2002.
XX
XX      30-NOV-2001; 2001WO-US047576.
XX
XX      12-DEC-2000; 2000US-00735705.
XX      07-MAY-2001; 2001US-00850716.
XX      28-JUN-2001; 2001US-00897778.
XX
XX      (CORI-) CORIXA CORP.
XX
XX      Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
XX      McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
XX      Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX
XX      WPI; 2002-583465/62.
XX
XX      Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
XX      the polynucleotides, useful in pharmaceutical compositions such as
XX      vaccines and as markers to indicate the presence of lung cancer.
XX
XX      Example 1; Page 313-314; 381pp; English.
XX
XX      The present invention describes isolated human lung carcinoma
XX      polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
XX      activity, and can be used in gene therapy and in vaccines. Compositions
XX      comprising (I) or (II) can be used for stimulating an immune response in
XX      a patient and for treating lung cancer in a patient. Oligonucleotides of
XX      (I) can be used for detecting the presence of a cancer in a patient, by
XX      obtaining a biological sample from the patient, contacting the biological
XX      sample with the oligonucleotide, detecting in the sample, an amount of
XX      polynucleotide that hybridises to the oligonucleotide and comparing the
XX      amount of polynucleotide that hybridises to the oligonucleotide to a
XX      predetermined cut-off value, and determining the presence of a cancer in
XX      the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
XX      vaccines. (I) is useful as a marker to indicate the presence or absence
XX      of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to
XX      ABP61992 represent sequences used in the exemplification of the present
XX      invention
XX
XX      Sequence 301 BP; 41 A; 109 C; 81 G; 67 T; 0 U; 3 Other;
XX
XX      Query Match 76.7%; Score 23; DB 6; Length 301;
XX      Best Local Similarity 95.8%; Pred. No. 23;
XX      Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0
XX
Qy      1  CGGAATAAAGGCTGTTGTAAGA 24
      |||
Db      25  CGNAATAAAGGCTGTTGTAAGA 2
|||||

Search completed: December 7, 2005, 00:03:14
Job time : 49.1624 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 13.094 Seconds
(without alignments)
4072.612 Million cell updates/sec

Title: US-09-980-046B-9

Perfect score: 30

Sequence: 1 cggaaataaaggctgtgtaagacaaaaa 30

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
- 9: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	23	76.7	301	3	US-09-643-597-304
C 2	23	76.7	301	3	US-09-480-884A-304
C 3	23	76.7	301	3	US-09-542-615A-304
C 4	23	76.7	301	3	US-09-606-421B-304
C 5	23	76.7	301	3	US-09-630-940B-304
C 6	23	76.7	301	3	US-10-007-700-304
C 7	21	70.0	634	3	US-09-917-254-33
C 8	20.6	68.7	245	3	US-09-270-767-85
C 9	20.6	68.7	245	3	US-09-270-767-15367
C 10	20.4	68.0	721	3	US-08-998-416-618
C 11	20.4	68.0	74790	3	US-09-949-016-15321
C 12	20	66.7	195	3	US-09-248-796A-14057
C 13	20	66.7	273	3	US-09-107-532A-1668
C 14	20	66.7	5220	2	US-08-777-405A-1
C 15	20	66.7	5220	2	US-08-977-871A-1
C 16	20	66.7	5220	2	US-09-225-951-1
C 17	20	66.7	5220	3	US-09-841-341-1
C 18	20	66.7	5220	3	US-10-027-591-1
C 19	20	66.7	5220	3	US-10-337-192-1
C 20	20	66.7	5220	3	US-10-697-912-1
C 21	20	66.7	1664976	3	US-08-916-421B-1
C 22	20	66.7	1664976	3	US-09-692-570-1
C 23	19.6	65.3	601	3	US-09-949-016-186914
C 24	19.6	65.3	123463	3	US-09-949-016-17078

25	19.4	64.7	601	3	US-09-949-016-60026	Sequence 60026, A
C 26	19.4	64.7	1065	3	US-09-540-236-488	Sequence 488, App
C 27	19.4	64.7	1108	3	US-09-838-561-14	Sequence 14, Appl
C 28	19.4	64.7	2742	3	US-09-134-000C-2635	Sequence 2635, Ap
C 29	19.4	64.7	48328	3	US-08-596-002-27	Sequence 27, Appl
C 30	19.4	64.7	343352	3	US-09-949-016-13498	Sequence 13498, A
C 31	19	63.3	450	3	US-09-235-451-11	Sequence 11, Appl
C 32	19	63.3	450	3	US-09-978-303-11	Sequence 20435, A
C 33	19	63.3	601	3	US-09-949-016-20435	Sequence 35678, A
C 34	19	63.3	601	3	US-09-949-016-35678	Sequence 35712, A
C 35	19	63.3	601	3	US-09-949-016-35712	Sequence 35746, A
C 36	19	63.3	601	3	US-09-949-016-35746	Sequence 47351, A
C 37	19	63.3	601	3	US-09-949-016-47351	Sequence 87631, A
C 38	19	63.3	601	3	US-09-949-016-87631	Sequence 87632, A
C 39	19	63.3	601	3	US-09-949-016-87632	Sequence 129248, A
C 40	19	63.3	601	3	US-09-949-016-129248	Sequence 129319, A
C 41	19	63.3	601	3	US-09-949-016-129319	Sequence 129390, A
C 42	19	63.3	601	3	US-09-949-016-129390	Sequence 129461, A
C 43	19	63.3	601	3	US-09-949-016-129461	Sequence 129532, A
C 44	19	63.3	601	3	US-09-949-016-129532	Sequence 129603, A
C 45	19	63.3	601	3	US-09-949-016-129603	

ALIGNMENTS

RESULT 1
US-09-643-597-304/c
; Sequence 304, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(301)
; OTHER INFORMATION: n = A,T,C or G
US-09-643-597-304
Query Match 76.7%; Score 23; DB 3; Length 301;
Best Local Similarity 95.8%; Pred. No. 4.1;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CGGAATAAAGGCTGTGTAAGA 24
|||
Db 25 CGNAATAAAGGCTGTGTAAGA 2
RESULT 2
US-09-480-884A-304/c
; Sequence 304, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(301)
; OTHER INFORMATION: n = A,T,C or G
US-09-480-884A-304

Query Match 76.7%; Score 23; DB 3; Length 301;
Best Local Similarity 95.8%; Pred. No. 4.1;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24
Db 25 CGNAATAAAGGCTGTTGTAAGA 2

RESULT 3

US-09-542-615A-304/c
; Sequence 304, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542.615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(301)
; OTHER INFORMATION: n = A,T,C or G
US-09-542-615A-304

Query Match 76.7%; Score 23; DB 3; Length 301;
Best Local Similarity 95.8%; Pred. No. 4.1;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24
Db 25 CGNAATAAAGGCTGTTGTAAGA 2

RESULT 4

US-09-606-421B-304/c
; Sequence 304, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(301)
; OTHER INFORMATION: n = A,T,C or G
US-09-606-421B-304

Query Match 76.7%; Score 23; DB 3; Length 301;
Best Local Similarity 95.8%; Pred. No. 4.1;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24
Db 25 CGNAATAAAGGCTGTTGTAAGA 2

RESULT 5

US-09-630-940B-304/c
; Sequence 304, Application US/09630940B
; Patent No. 6737514
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C10
; CURRENT APPLICATION NUMBER: US/09/630,940B
; CURRENT FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(301)
; OTHER INFORMATION: n = A,T,C or G
US-09-630-940B-304

Query Match 76.7%; Score 23; DB 3; Length 301;
Best Local Similarity 95.8%; Pred. No. 4.1;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24

Db 25 CGNAATAAAGGCTGTTGTAAGA 2
|| |||||

RESULT 6
US-10-007-700-304/c
; Sequence 304, Application US/10007700
; Patent No. 6960570
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007.700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 23, 104, 192
; OTHER INFORMATION: n = A,T,C or G
US-10-007-700-304

Query Match 76.7%; Score 23; DB 3; Length 301;
Best Local Similarity 95.8%; Pred. No. 4.1;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTAAGA 24
|| |||||

Db 25 CGNAATAAAGGCTGTTGTAAGA 2
|| |||||

RESULT 7
US-09-917-254-33
; Sequence 33, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
; APPLICANT: Mutter, George
; APPLICANT: Baak, Jan
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
; FILE REFERENCE: B0801/7224(JRV)
; CURRENT APPLICATION NUMBER: US/09/917,254
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/222,093
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-917-254-33

Query Match 70.0%; Score 21; DB 3; Length 634;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGAATAAAGGCTGTTGTA 21
|| |||||

Db 608 CGGAATAAAGGCTGTTGTA 628
|| |||||

RESULT 8
US-09-270-767-85/c
; Sequence 85, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 245
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-85

Query Match 68.7%; Score 20.6; DB 3; Length 245;
Best Local Similarity 85.2%; Pred. No. 37;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 AAATAAAGGCTGTTGTAAGACAAAA 30
|| |||||

Db 237 AAATAAAGGCTGGGTACACAAAA 211
|| |||||

RESULT 9
US-09-270-767-15367/c
; Sequence 15367, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15367
; LENGTH: 245
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-15367

Query Match 68.7%; Score 20.6; DB 3; Length 245;
Best Local Similarity 85.2%; Pred. No. 37;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 AAATAAAGGCTGTTGTAAGACAAAA 30
|| |||||

Db 237 AAATAAAGGCTGGGTACACAAAA 211
|| |||||

RESULT 10
US-08-998-416-618
; Sequence 618, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Fohmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen


```
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1668:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...273
; SEQUENCE DESCRIPTION: SEQ ID NO: 1668:
US-09-107-532A-1668

Query Match 66.7%; Score 20; DB 3; Length 273;
Best Local Similarity 82.1%; Pred. No. 66;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GGAATAAAGCGCTGTGTAAAGACAAAA 29
   ||||| ||||| ||||| ||||| |||||
Db 8 GGAAGAAGACGCTGTGTAAAGACACA 35

RESULT 14
US-08-777-405A-1
; Sequence 1, Application US/0877405A
; Patent No. 5858753
; GENERAL INFORMATION:
; APPLICANT: Chantry, David
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: No. 5858753el Lipid Kinase
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall O'Toole Gerstein Murray & Borun
; STREET: 6300 Sears Tower/233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/777,405
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5858753and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/33441
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5220 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 196..3327
US-08-977-871A-1

Query Match 66.7%; Score 20; DB 2; Length 5220;
Best Local Similarity 82.1%; Pred. No. 1e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 196..3327
US-08-777-405A-1

Query Match 66.7%; Score 20; DB 2; Length 5220;
Best Local Similarity 82.1%; Pred. No. 1e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 GAAATAAAGCGCTGTGTAAAGACAAAA 30
   ||||| ||||| ||||| ||||| |||||
Db 5185 GAAGTAAGGCAGATGAAGAGAAAAAA 5212

RESULT 15
US-08-977-871A-1
; Sequence 1, Application US/08977871A
; Patent No. 5882910
; GENERAL INFORMATION:
; APPLICANT: Chantry, David
; APPLICANT: Hoekstra, Merl F.
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: No. 5882910el Lipid Kinase
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall O'Toole Gerstein Murray & Borun
; STREET: 6300 Sears Tower/233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/977,871A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/777,405
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5882910and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/33441
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5220 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 196..3327
US-08-977-871A-1

Query Match 66.7%; Score 20; DB 2; Length 5220;
Best Local Similarity 82.1%; Pred. No. 1e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 GAAATAAAGCGCTGTGTAAAGACAAAA 30
   ||||| ||||| ||||| ||||| |||||
Db 5185 GAAGTAAGGCAGATGAAGAGAAAAAA 5212

Search completed: December 6, 2005, 11:47:28
```

Job time : 15.094 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:34:11 ; Search time 20.4786 Seconds
(without alignments)
455.919 Million cell updates/sec

Title: US-09-980-046B-9
Perfect score: 30
Sequence: 1 cggaaataaaggctgttgaagacaaaaa 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New:*
1: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
2: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
5: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubna/US11_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubna/US11_NEW_PUB.seq.*
9: /cgn2_6/ptodata/1/pubna/US11_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	23	76.7	301	6	US-10-623-155-304
C 2	21.2	70.7	940	6	US-10-750-185-47224
C 3	20	66.7	600	6	US-10-750-185-151
C 4	20	66.7	1583	6	US-10-750-185-27698
C 5	20	66.7	2175	6	US-10-750-185-60060
C 6	20	66.7	5220	7	US-11-110-204-1
C 7	19.4	64.7	811	6	US-10-750-185-41266
C 8	19.4	64.7	1490	6	US-10-750-185-42248
C 9	19.4	64.7	1530	6	US-10-750-185-64471
C 10	19.4	64.7	1810	6	US-10-750-185-57881
C 11	19	63.3	19	8	US-11-101-244-663428
C 12	19	63.3	19	8	US-11-101-244-663430
C 13	19	63.3	19	9	US-11-083-784-663430
C 14	19	63.3	19	9	US-11-083-784-663430
C 15	18.8	62.7	1290	6	US-10-467-657-661
C 16	18.8	62.7	1290	6	US-10-467-657-4925
C 17	18.8	62.7	1632	6	US-10-750-185-38307
C 18	18.8	62.7	1956	6	US-10-750-185-45040
C 19	18.8	62.7	4047	6	US-10-750-185-61814
C 20	18.6	62.0	600	6	US-10-750-185-21234
C 21	18.6	62.0	892	6	US-10-750-185-55928
C 22	18.6	62.0	1160	6	US-10-750-185-31088
C 23	18.6	62.0	1305	6	US-10-750-185-51810

C 24	18.4	61.3	1050	6	US-10-793-626-1449	Sequence 1449, Ap
C 25	18.4	61.3	1084	6	US-10-750-185-49927	Sequence 49927, A
C 26	18.4	61.3	1112	6	US-10-750-185-54938	Sequence 54938, A
C 27	18.4	61.3	1113	6	US-10-793-626-1115	Sequence 1115, Ap
C 28	18.4	61.3	1113	6	US-10-793-626-2131	Sequence 2131, Ap
C 29	18.4	61.3	1259	6	US-10-750-185-59326	Sequence 59326, A
C 30	18.4	61.3	1259	6	US-10-793-626-1909	Sequence 1909, Ap
C 31	18.4	61.3	3037	6	US-10-793-626-3904	Sequence 3904, Ap
C 32	18.4	61.3	3180	6	US-10-793-626-3665	Sequence 3665, Ap
C 33	18.4	61.3	3394	6	US-10-793-626-3365	Sequence 3365, Ap
C 34	18.4	61.3	3536	6	US-10-793-626-3957	Sequence 3957, Ap
C 35	18.2	60.7	1987	6	US-10-750-185-25087	Sequence 25087, A
C 36	18	60.0	19	8	US-11-101-244-663437	Sequence 663437, A
C 37	18	60.0	19	9	US-11-083-784-663437	Sequence 45289, A
C 38	18	60.0	1037	6	US-10-750-185-45289	Sequence 45289, A
C 39	18	60.0	1900	6	US-10-750-185-63880	Sequence 63880, A
C 40	18	60.0	1952	6	US-10-750-185-51869	Sequence 51869, A
C 41	18	60.0	2072	6	US-10-750-185-52444	Sequence 52444, A
C 42	18	60.0	2144	6	US-10-750-185-42852	Sequence 42852, A
C 43	18	60.0	2382	6	US-10-750-185-25473	Sequence 25473, A
C 44	18	60.0	150468	7	US-11-112-908-56	Sequence 56, Appl
C 45	18	60.0	193789	7	US-11-112-908-55	Sequence 55, Appl

ALIGNMENTS

RESULT 1
US-10-623-155-304/c
; Sequence 304, Application US/10623155
; Publication No. US20050261166A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Peckham, David W.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C20
; CURRENT APPLICATION NUMBER: US/10/623.155
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 304
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 23, 104, 192
; OTHER INFORMATION: n = A,T,C or G

Query Match 76.7%; Score 23; DB 6; Length 301;
Best Local Similarity 95.8%; Pred. No. 1.1;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGAATAAAGCCTGTTGTAAGA 24
|||
DB 25 CGNAATAAAGCCTGTTGTAAGA 2

RESULT 2
US-10-750-185-47224/c
; Sequence 47224, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: Denise, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen

```
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47224
; LENGTH: 940
; TYPE: DNA
; ORGANISM: Bovine 19866881078624
US-10-750-185-47224

Query Match      70.7%; Score 21.2; DB 6; Length 940;
Best Local Similarity 88.5%; Pred. No. 6.6;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      4 AAATAAAGGCTGTGTTAAAGACAAAA 29
Db      323 AACTTAAGGCTGTGTTGAAGACAAAA 298

RESULT 3
US-10-750-185-151
; Sequence 151, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 151
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT02951
US-10-750-185-151

Query Match      66.7%; Score 20; DB 6; Length 600;
Best Local Similarity 82.1%; Pred. No. 18;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      2 GGAATAAAGGCTGTGTTAAAGACAAAA 29
Db      118 GGTGATAAGGCTATTGAAAGACAAAA 145

RESULT 4
US-10-750-185-27698
; Sequence 27698, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 151
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT02951
US-10-750-185-151

Query Match      66.7%; Score 20; DB 6; Length 600;
Best Local Similarity 82.1%; Pred. No. 18;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      2 GGAATAAAGGCTGTGTTAAAGACAAAA 29
Db      118 GGTGATAAGGCTATTGAAAGACAAAA 145

RESULT 5
US-10-750-185-60060
; Sequence 60060, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60060
; LENGTH: 2175
; TYPE: DNA
; ORGANISM: Bovine 19866880609486
US-10-750-185-60060

Query Match      66.7%; Score 20; DB 6; Length 2175;
Best Local Similarity 82.1%; Pred. No. 23;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      3 GGAATAAAGGCTGTGTTAAAGACAAAA 30
Db      1428 GAAAAAATACTGTTGTGAAACAAAAA 1455

RESULT 6
US-11-110-204-1
; Sequence 1, Application US/11110204
; Publication No. US20050261317A1
; GENERAL INFORMATION:
; APPLICANT: SADHU, Chanchal et al.
; TITLE OF INVENTION: INHIBITORS OF HUMAN PHOSPHATIDYLINOSITOL 3-KINASE DELTA
; FILE REFERENCE: 27866/361708
; CURRENT APPLICATION NUMBER: US/11/110,204
; CURRENT FILING DATE: 2005-04-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 5220
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
```

```
; OTHER INFORMATION: Human pilodelta complete cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (196)...(3327)
US-11-110-204-1

Query Match      66.7%; Score 20; DB 7; Length 5220;
Best Local Similarity 82.1%; Pred. No. 27;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy  3 GAAATAAAGGCTGTTGTAAGACAAAAA 30
    ||| ||||| ||| ||||| ||||| |||||
Db  5185 GAAGTAAGGCAGATGAAGAAAAA 5212

RESULT 7
US-10-750-185-41266/c
; Sequence 41266, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 41266
; LENGTH: 811
; TYPE: DNA
; ORGANISM: Bovine 19866880536641
US-10-750-185-41266

Query Match      64.7%; Score 19.4; DB 6; Length 811;
Best Local Similarity 79.3%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy  2 GGAATAAAGGCTGTTGTAAGACAAAAA 30
    ||| ||||| ||| ||||| ||||| |||||
Db  789 GGTAAATATCAGTCTGTTAAGACAAATA 761

RESULT 8
US-10-750-185-42248/c
; Sequence 42248, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 42248
; LENGTH: 1490
; TYPE: DNA
; ORGANISM: Bovine 19866880528816
US-10-750-185-42248

; ORGANISM: Bovine 19866880695859
US-10-750-185-42248

Query Match      64.7%; Score 19.4; DB 6; Length 1490;
Best Local Similarity 79.3%; Pred. No. 37;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy  2 GGAATAAAGGCTGTTGTAAGACAAAAA 30
    ||| ||||| ||| ||||| ||||| |||||
Db  872 GGAAGAATGCTGGGTAAAAA 844

RESULT 9
US-10-750-185-64471/c
; Sequence 64471, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 64471
; LENGTH: 1530
; TYPE: DNA
; ORGANISM: Bovine 19866880491953
US-10-750-185-64471

Query Match      64.7%; Score 19.4; DB 6; Length 1530;
Best Local Similarity 79.3%; Pred. No. 37;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy  2 GGAATAAAGGCTGTTGTAAGACAAAAA 30
    ||| ||||| ||| ||||| ||||| |||||
Db  125 GGAATAAAGATGGTTTAAAAA 97

RESULT 10
US-10-750-185-57881
; Sequence 57881, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 57881
; LENGTH: 1810
; TYPE: DNA
; ORGANISM: Bovine 19866880528816
US-10-750-185-57881
```

Query Match 64.7%; Score 19.4; DB 6; Length 1810;
Best Local Similarity 79.3%; Pred. No. 38;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 GGAATAAAGGCTGTTGTTAAAGACAAAA 30
|||||
Db 1183 GGAATAAAGACTATCTTTAAAAA 1211

```

RESULT 11
US-11-101-244-663428
; Sequence 663428, Application US/11101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCES: 13495US
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 663428
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-663428

```

Query Match 63.3%; Score 19; DB 8; Length 19;
Best Local Similarity 73.7%; Pred. NO. 24;
Matches 14; Conservative 5; Mismatches 0; Indels

Qy 2 GGAAATAAAGGCTGTTGTA 20
|||||:|||||:|:
Db 1 GGAAAUAAAGGCUGUUGUA 19

```

RESULT 12
US-11-101-244-663430
; Sequence 663430, Application US/11101244
; Publication No. US20050246794A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 1349SUS
; CURRENT APPLICATION NUMBER: US/11/101,244
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 663430
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-101-244-663430

```

Query Match 63.3%; Score 19; DB 8; Length 19;

Qy	3	GAATAAAGCCTCTTGTAA	21
		: : : :	
		: : : :	
Db	1	GAATAAAGCGUGTUGAA	19

Best Local Similarity 73.7%; Pred. No. 24;
 Matches 14; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 13
US-11-083-784-663428
; Sequence 663428, Application US/11083784
; Publication No. US20050245475A1
; GENERAL INFORMATION:
; APPLICANT: Dharmacon, Inc.
; APPLICANT: Khvorova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scarsange, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 13499US
; CURRENT APPLICATION NUMBER: US/11/083,784
; CURRENT FILING DATE: 2005-03-18
; PRIOR APPLICATION NUMBER: US/10/714,333
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 663428
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-11-083-784-663428

```

Query Match	63.3%;	Score 19;	DB 9;	Length 19;
Best Local Similarity	73.7%;	Pred. No. 24;		
Matches 14;	Conservative	5;	Mismatches	0;
			Indels	0;
			Gaps	0;

Qy 2 GGAAATAAAGGCTGTGTGA 20
|||:|||||:|:
Db 1 GGAAAUAAAGGCUGUUGUA 19

```

RESULT 14
US-11-083-784-663430
; Sequence 663430, Application US/11083784
; Publication No. US20050245475A1
; GENERAL INFORMATION:
; APPLICANT: Dharmakon, Inc.
; APPLICANT: Khvorotova, Anastasia
; APPLICANT: Reynolds, Angela
; APPLICANT: Leake, Devin
; APPLICANT: Marshall, William
; APPLICANT: Scaringe, Stephen
; TITLE OF INVENTION: Functional and Hyperfunctional siRNA
; FILE REFERENCE: 134990S
; CURRENT APPLICATION NUMBER: US/11/083,784
; CURRENT FILING DATE: 2005-03-18
; PRIORITY APPLICATION NUMBER: US/10/714,333
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 60/502,050
; PRIOR FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/426,137
; PRIOR FILING DATE: 2002-11-14
; NUMBER OF SEQ ID NOS: 1591911
; SOFTWARE: Proprietary
; SEQ ID NO 663430
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens

```

```

;
; TYPE: RNA
; ORGANISM: Homo sapiens

```


US-11-083-784-663430

Query Match 63.3%; Score 19; DB 9; Length 19;
 Best Local Similarity 73.7%; Pred. No. 24;
 Matches 14; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GAAATAAAGGCTGTGTAA 21
 ||||:|||||:|:|:
 Db 1 GAAAUAAAGGCGUGUAA 19

RESULT 15

US-10-467-657-661
 ; Sequence 661, Application US/10467657
 ; Publication No. US20050260581A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON Spa
 ; APPLICANT: FONTANA Maria Rita
 ; APPLICANT: PIZZA Mariagrazia
 ; APPLICANT: MASIGNANI Vega
 ; APPLICANT: MONACI Elisabetta
 ; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
 ; FILE REFERENCE:
 ; CURRENT APPLICATION NUMBER: US/10/467,657
 ; CURRENT FILING DATE: 2003-08-11
 ; PRIOR APPLICATION NUMBER: GB-0103424.8
 ; PRIOR FILING DATE: 2001-02-12
 ; NUMBER OF SEQ ID NOS: 9218
 ; SOFTWARE: SeqWin99, version 1.04
 ; SEQ ID NO 661
 ; LENGTH: 1290
 ; TYPE: DNA
 ; ORGANISM: Neisseria gonorrhoeae
 US-10-467-657-661

Query Match 62.7%; Score 18.8; DB 6; Length 1290;
 Best Local Similarity 90.9%; Pred. No. 62;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GGAATAAAGGCTGTGTAAAG 23
 |||||:|||||:|:|:
 Db 648 GGAATACAGGCTGTCGTAAAG 669

Search completed: December 6, 2005, 23:16:55
 Job time : 21.4786 secs

This Page Blank (uspto)

This Page Blank (uspto)

This Page Blank (uspto)

This Page Blank (uspto)

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 11:47:47 ; Search time 133.627 Seconds
(without alignments)
4239.411 Million cell updates/sec

Title: US-09-980-046B-10

Perfect score: 85

Sequence: 1 cggtagagtcacatcttcgc.....tgccccactaagtagaaaaa 85

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*
- 14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	100.0	85	4	AAC89741 Human gas
2	83.4	98.1	451	4	AAS04644 Gene expr
3	83.4	98.1	1126	13	AQ38498 Human SNP
4	83.4	98.1	1143	3	AAF18060 Lung can
5	83.4	98.1	1271	10	ADJ56466 Human CDN
6	83.4	98.1	1458	13	AQD83779 Human tum
7	83.4	98.1	1539	13	AQD38499 Human SNP
8	83.4	98.1	2027	11	ACN89244 Breast ca
9	83.4	98.1	3674	10	ADJ56452 Human CDN
10	83.4	98.1	66109	6	ABL62723 Colon ade
11	83.4	98.1	66109	6	ABN97273 Gene #377
12	83.4	98.1	198285	6	ABK84699 Human CDN
13	83.4	98.1	198285	6	ABN97319 Gene #381
14	83.4	98.1	198285	13	ADR52987 Drug ther
15	81.8	96.2	108	4	AAC89745 Human gas
16	54.6	64.2	11398	6	ABL34438 Human imm
17	54.6	64.2	11398	6	ABL34439 Human imm
18	28	32.9	1205	10	ADF76527 Novel hum
19	28	32.9	1205	13	ACN38440 Tumour-as

c	20	28	32.9	2379	8	ADA69895
c	21	28	32.9	2379	10	ADC07941
c	22	27.4	32.2	44770	11	ACN43900
c	23	27.2	32.0	474	13	ACN47800
c	24	27	31.8	2489	4	AAI64599
c	25	26.8	31.5	1002	2	AAX85552
c	26	26.8	31.5	1296	2	AAX85549
c	27	26.8	31.5	1983	4	ABL21789
c	28	26.8	31.5	4097	4	ABL21788
c	29	26.4	31.1	31449	4	ABL09172
c	30	26	30.6	336	14	ADY21257
c	31	26	30.6	529	12	ADP64404
c	32	26	30.6	716	12	ADP64412
c	33	26	30.6	856	14	ADW95072
c	34	26	30.6	857	14	ADW95077
c	35	26	30.6	2000	6	ABZ15407
c	36	26	30.6	3768	8	ACA27605
c	37	25.8	30.4	1390	2	AAX85570
c	38	25.8	30.4	1458	2	AAQ79936
c	39	25.6	30.1	323	3	AAC26084
c	40	25.6	30.1	376	11	ADT96789
c	41	25.6	30.1	376	11	ADX43271
c	42	25.6	30.1	637	11	ADT94682
c	43	25.6	30.1	637	11	ADX41164
c	44	25.6	30.1	793	8	ABZ52278
c	45	25.6	30.1	3589	10	ADF12420

ALIGNMENTS

RESULT 1

AAC89741

ID AAC89741 standard; cDNA; 85 BP.

XX AAC89741;

XX 12-MAR-2001 (first entry)

XX Human gastrointestinal inflammation-related CDNA, SEQ ID NO: 10.

XX Human; cytostatic; immunomodulator; immunostimulant; vulneryary;
XX anti-inflammatory; neuroprotective; antibacterial; gene therapy;
XX gastrointestinal inflammation; immune system disorder; genetic disorder;
XX cancer; autoimmune disorder; infection; wound healing; ss.

XX Homo sapiens.

XX WO200073324-A2.

XX 07-DEC-2000.

XX 01-JUN-2000; 2000WO-US015191.

XX 01-JUN-1999; 99US-0137058P.

(DIGI-) DIGITAL GENE TECHNOLOGIES INC.

XX Youakim A, Dubose RF, Sims JB, Pribyl TM, Hillbush BS, Hasel KW;

XX WPI; 2001-061508/07.

XX New polynucleotides and polypeptides, useful in gene therapy and in
XX diagnosing a pathological condition, e.g. for modulating gene expression
XX in gastrointestinal inflammation, or for treating cancers or genetic
XX disorders.

XX Claim 1; Page 86; 108pp; English.

XX The present sequence is one of a number of isolated human polynucleotides
XX which are useful in gene therapy, and for diagnosing a pathological
XX condition or a susceptibility to it. In particular, the polynucleotides
XX are useful for modulating gene expression in gastrointestinal

CC inflammation. The polynucleotides are useful for chromosome
CC identification, controlling gene expression through triple helix
CC formation or antisense DNA or RNA, or identifying individuals from minute
CC biological samples using DNA-based identification techniques. The
CC polynucleotides can also be used as an alternative to restriction
CC fragment length polymorphism (RFLP), by determining the actual base-by-
CC base DNA sequences of selected portions of an individual's genome. The
CC polynucleotides may also be used as molecular weight markers on Southern
CC gels, as diagnostic probes for the presence of a specific mRNA, as a
CC probe to subtract-out known sequences in the process of discovering novel
CC polynucleotides, or as an antigen to elicit an immune response. The
CC polypeptides are useful in diagnostic procedures to detect a disorder.
CC The polynucleotides and polypeptides are useful for preventing, treating
CC or ameliorating immune system disorders, genetic disorders, cancers, some
CC autoimmune disorders, or infections. The polynucleotides and polypeptides
CC are also useful for differentiating, proliferating or attracting cells,
CC leading to the regeneration of tissues, especially in wounds or burns.
CC The polypeptides and polynucleotides may also be used as a food additive
CC or preservative
XX
SQ Sequence 85 BP; 21 A; 20 C; 20 G; 24 T; 0 U; 0 Other;

Query Match 100.0%; Score 85; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 2.2e-21;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT 60
DB 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT 60
QY 61 CTGGGTCCCCCACTAAGTAGAAAA 85
DB 61 CTGGGTCCCCCACTAAGTAGAAAA 85

RESULT 2
AAS04644
ID AAS04644 standard; cDNA; 451 BP.
XX
AC AAS04644;
DT
DT 07-SEP-2001 (first entry)
XX
XX Gene expression profile sequence #144.
XX
XX Gene expression profile; hypersensitivity; DNA microarray;
KW liver toxicity; hepatitis; tumour formation; immunosuppression;
KW renal toxicity; glomerulitis; neurotoxicity; leukaemia; dementia;
KW peripheral neuropathy; hypertension; hypotension; myelosuppression;
KW retinopathy; inflammation; sensitisation; ss.
XX
OS Homo sapiens.
XX
XX WO200132928-A2.
XX
XX 10-MAY-2001.
XX
XX 03-NOV-2000; 2000WO-US030474.
XX
XX 05-NOV-1999; 99US-01653398P.
XX
XX 11-APR-2000; 2000US-0196571P.
XX
XX (PHAS-) PHASE-1 MOLECULAR TOXICOLOGY.
XX
XX Farr S;
XX
XX WPI; 2001-328806/34.
XX
XX
XX Identifying hypersensitivity in a subject by obtaining a gene expression
PT profile of hypersensitivity associated genes and detecting a
PT predetermined pattern of gene expression of hypersensitivity associated
PT genes.
XX

PS Claim 24; Page 176; 222pp; English.
XX
CC The sequence represents a cDNA from a gene associated with
CC hypersensitivity to a agent, the sequence was detected in a sample by use
CC of a DNA microarray containing genes from a gene expression profile
CC thought to be associated with hypersensitivity to an agent. The invention
CC relates to methods of obtaining a gene expression profile of genes
CC associated with hypersensitivity to an agent involving comparing the gene
CC expression profile of cells treated with the agent with the gene
CC expression profile of cells not treated with the agent, and determining
CC the genes that have altered expression due to exposure to the agent.
CC Hypersensitivity in a subject can then be detected by comparing the gene
CC expression profile of the subject with that associated with the
CC hypersensitivity, usually by hybridisation of a sample of mRNA or cDNA
CC from the subject to a DNA microarray containing genes from the
CC hypersensitivity profile. The genes in the profiles are associated with
CC liver toxicity (e.g. hepatitis), tumour formation, immunosuppression,
CC renal toxicity (e.g. glomerulitis), neurotoxicity, leukaemia, dementia,
CC peripheral neuropathy, hyper/hypotension, myelosuppression, retinopathy,
CC inflammation, and sensitisation
XX
SQ Sequence 451 BP; 119 A; 100 C; 135 G; 97 T; 0 U; 0 Other;
Query Match 98.1%; Score 83.4; DB 4; Length 451;
Best Local Similarity 98.8%; Pred. No. 1.5e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT 60
DB 353 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCCAGGGAAGTTT 412
QY 61 CTGGGTCCCCCACTAAGTAGAAAA 85
DB 413 CTGGGTCCCCCACTAAGTAGATAA 437
RESULT 3
ADQ38498
ID ADQ38498 standard; DNA; 1126 BP.
XX
AC ADQ38498;
XX
XX 18-NOV-2004 (first entry)
DT
XX
DE Human SNP containing myocardial infarction-associated gene, SEQ ID 161.
XX
XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
XX WO2004058052-A2.
XX
XX 15-JUL-2004.
XX
XX 22-DEC-2003; 2003WO-US040978.
XX
XX 20-DEC-2002; 2002US-0434778P.
PR 10-MAR-2003; 2003US-0453135P.
PR 30-APR-2003; 2003US-0466412P.
PR 23-SEP-2003; 2003US-0504955P.
XX
XX (APPL-) APPLERA CORP.
PA
XX Cargill M, Devlin JJ, Takubova O;
PI
XX WPI; 2004-533949/51.
XX
XX P-PSDB; ADQ39326.
DR
XX
XX Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX

PS Claim 7; SEQ ID NO 161; 145pp; English.

XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a human myocardial infarction-
CC associated gene containing one or more SNP's of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.

XX SQ Sequence 1126 BP; 259 A; 279 C; 339 G; 247 T; 0 U; 2 Other;

Query Match 98.1%; Score 83.4; DB 13; Length 1126;

Best Local Similarity 98.8%; Pred. No. 2.1e-20;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60

Db 1022 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1081

Qy 61 CTGGTGCCCCCACTAAGTAGAAAAA 85

Db 1082 CTGGTGCCCCCACTAAGTAGAATAA 1106

RESULT 4

AAFI8060

ID AAFI8060 standard; DNA; 1143 BP.

XX

AC AAFI8060;

XX 14-MAR-2001 (first entry)

DT

XX Lung cancer associated polynucleotide sequence SEQ ID 79.

DE

XX Human; lung cancer associated protein; neuroprotective; cytostatic;

KW cardioactive; immunomodulatory; muscular active; vulnerary;

KW gastrointestinal; nephrotropic; antiinfective; gynecological;

KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;

KW proliferative disorder; wound healing; infectious disease; ds.

XX

OS Homo sapiens.

XX

PN WO200055180-A2.

XX

PD 21-SEP-2000.

XX

PF 08-MAR-2000; 2000WO-US005918.

XX

PR 12-MAR-1999; 99US-0124270P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

XX Ruben SM;

PI

XX WPI; 2000-587514/55.

DR P-PSDB; AAB58184.

XX

PT Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer.

XX

PS Claim 1; Page 556; 1425pp; English.

XX

CC Polynucleotide sequences AAFI7982 - AAFI8424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytostatic; cardioactive;
CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
CC general; nephrotropic; antiinfective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the protein
CC or polynucleotide sequences. The lung cancer associated polynucleotide
CC sequences may be used for detection of lung cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The proteins may be used to treat disorders such as
CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary, also
CC cardiovascular, renal, and proliferative disorders. The proteins may also
CC be used in the treatment of wounds and infectious diseases.

CC Polynucleotide sequences AAFI8425 - AAFI8433 and peptide AAB58549 are
CC used in the course of the invention for the identification and
CC characterisation of the polynucleotide and protein sequences

XX

SQ Sequence 1143 BP; 276 A; 275 C; 348 G; 242 T; 0 U; 2 Other;

Query Match 98.1%; Score 83.4; DB 3; Length 1143;

Best Local Similarity 98.8%; Pred. No. 2.1e-20;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60

Db 1006 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1065

Qy 61 CTGGTGCCCCCACTAAGTAGAAAAA 85

Db 1066 CTGGTGCCCCCACTAAGTAGAATAA 1090

RESULT 5

ADJ56466

ID ADJ56466 standard; cDNA; 1271 BP.

XX

AC ADJ56466;

XX

DT 06-MAY-2004 (first entry)

XX

DE Human cDNA differentially expressed in MYCN activated cells SeqID 272.

XX

XX human; differential expression; transactivator; proto-oncogene;

KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;

KW MYCN activated cell.

XX

OS Homo sapiens.

XX

PN US2003119009-A1.

XX

PD 26-JUN-2003.

XX

PF 25-FEB-2002; 2002US-00084817.

XX

PR 23-FEB-2001; 2001US-0270784P.

XX

PA (STUA/) STUART S G.

PA (NUCH/) NUCHTERN J G.

PA (PLON/) PLON S E.

PA (SHOH/) SHOHET J M.

XX Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
XX WPI; 2003-635698/60.
XX New genes regulated by MYCN activation, useful in gene therapy,
XX particularly for treating a subject with e.g. neuroblastoma or other
XX cancers, or for diagnosing, staging or monitoring the treatment of the
XX cancer.
XX Claim 1; SEQ ID NO 272; 27pp; English.
XX This invention relates to novel isolated cDNAs that are differentially
XX expressed in MYCN activated cells. Specifically, it refers to
XX polynucleotide sequences that exhibit differential expression patterns in
XX cells activated by the transactivator MYCN, where MYCN is a proto-
XX oncogene that is amplified in neuroblastoma cells and is common in small
XX cell lung cancers. The present invention describes these cDNA molecules
XX as useful for in hybridisation assays to detect expression of nucleic
XX acids (or complementary nucleic acids) in a present in a given sample, as
XX well as for screening assays by identifying molecules or compounds that
XX specifically bind the cDNA as a ligand and modulate function or activity.
XX Accordingly, these compositions exhibit cytostatic activity and can also
XX be used for gene therapy purposes. This polynucleotide sequence is a cDNA
XX that is differentially expressed in MYCN activated cells, given in an
XX exemplification of the invention. NOTE: This sequence does not appear in
XX the printed specification but has been obtained in electronic format from
XX the US Patent Office at
XX ftp.seqdata.uspto.gov/sequence.html?DocID=20030119009.
XX Sequence 1271 BP; 295 A; 317 C; 379 G; 280 T; 0 U; 0 Other;
SQ
Query Match 98.1%; Score 83.4; DB 10; Length 1271;
Best Local Similarity 98.8%; Pred. No. 2.2e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1160 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1219
QY 61 CTGGTGCCCCCACTAAGTAGAATAA 85
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1220 CTGGTGCCCCCACTAAGTAGAATAA 1244
RESULT 6
ADQ83779
ID ADQ83779 standard; cDNA; 1458 BP.
XX
AC ADQ83779;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #593.
XX
XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
XX cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
XX W02004060270-A2.
XX
XX 22-JUL-2004.
XX
XX 15-OCT-2003; 2003WO-US029126.
XX
XX 18-OCT-2002; 2002US-0418988P.
XX
XX (GETH) GENENTECH INC.
XX
XX (WUTD/) WU T D.
XX
XX (ZHOU/) ZHOU Y.
XX
XX Wu TD, Zhou Y;

XX WPI; 2004-534300/51.
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
XX preventing or treating cell proliferative disorders such as cancer.
XX Claim 1; SEQ ID NO 593; 5504pp; English.
XX The present invention describes an isolated tumour-associated antigenic
XX target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
XX sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
XX (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
XX sequence identity to (a) - (c); or (e) a sequence that hybridises to (a) -
XX (c). Also described: (1) an expression vector comprising the above
XX nucleic acid; (2) a host cell comprising the above expression vector; (3)
XX a process for producing a polypeptide; (4) an isolated polypeptide
XX comprising: (a) an amino acid sequence encoded by any of the above
XX nucleotide sequences; (b) an amino acid sequence encoded by the full-
XX length coding region of the above nucleotide sequences; or (c) a sequence
XX having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
XX comprising the above polypeptide fused to a heterologous polypeptide; (6)
XX an isolated antibody that binds to the above polypeptide; (7) a process
XX for producing the antibody; (8) an isolated oligopeptide that binds to
XX the above polypeptide; (9) a tumour-associated antigenic target (TAT)
XX binding organic molecule that binds to the above polypeptide; (10) a
XX composition of matter comprising the above (chimeric) polypeptide,
XX antibody, oligopeptide or TAT binding organic molecule, in combination
XX with a carrier; (11) an article of manufacture comprising a container and
XX the composition of matter contained within the container; (12) methods of
XX inhibiting the growth of a cell that expresses the above protein, where
XX the growth of the cell is at least in part dependent upon a growth
XX potentiating effect of the above protein; (13) a method of
XX therapeutically treating a mammal having a cancerous tumour comprising
XX cells that express the above protein; (14) a method of determining the
XX presence of a protein in a sample suspected of containing the protein
XX described above; (15) methods of diagnosing the presence of a tumour in a
XX mammal; (16) a method for treating or preventing a cell proliferative
XX disorder associated with increased expression or activity of the above
XX protein; and (17) a method of binding an antibody, oligopeptide or
XX organic molecule to a cell that expresses the protein described above.
XX The TAT sequences have cytostatic activities, and can be used in gene
XX therapy. The composition and methods are useful for diagnosing,
XX preventing or treating cancer. The composition is also used for preparing
XX a medicament for the therapeutic treatment or diagnostic detection of a
XX cell proliferative disorder or cancer. The present sequence represents a
XX human TAT cDNA sequence from the present invention.
SQ Sequence 1458 BP; 331 A; 378 C; 421 G; 328 T; 0 U; 0 Other;
Query Match 98.1%; Score 83.4; DB 13; Length 1458;
Best Local Similarity 98.8%; Pred. No. 2.3e-20;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1355 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1414
QY 61 CTGGTGCCCCCACTAAGTAGAATAA 85
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1415 CTGGTGCCCCCACTAAGTAGAATAA 1439
RESULT 7
ADQ38499
ID ADQ38499 standard; DNA; 1539 BP.
XX
XX ADQ38499;
XX
XX 18-NOV-2004 (first entry)
XX
XX Human SNP containing myocardial infarction-associated gene, SEQ ID 162.
XX
XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;

KW cardiant; gene therapy; human; gene; ds.

XX Homo sapiens.

XX WO2004058052-A2.

XX 15-JUL-2004.

XX 22-DEC-2003; 2003WO-US040978.

XX 20-DEC-2002; 2002US-0434778P.

XX 10-MAR-2003; 2003US-0453135P.

XX 30-APR-2003; 2003US-0466412P.

XX 23-SEP-2003; 2003US-0504955P.

XX (APPL-) APPLERA CORP.

XX Cargill M, Devlin JJ, Iakoubova O;

XX WPI; 2004-533949/51.

XX P-PSDB; ABQ39327.

XX Identifying an individual who has an altered risk for developing

XX myocardial infarction by detecting a single nucleotide polymorphism in

XX the individual's nucleic acids.

XX Claim 7; SEQ ID NO 162; 145pp; English.

XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a human myocardial infarction-
CC associated gene containing one or more SNP's of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.

XX Sequence 1539 BP; 346 A; 405 C; 454 G; 332 T; 0 U; 2 Other;

Query Match 98.1%; Score 83.4; DB 13; Length 1539;

Best Local Similarity 98.8%; Pred. No. 2.3e-20;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60

Db 1435 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1494

Qy 61 CTGGGTGCCCACTAAGTAGAAAA 85

Db 1495 CTGGGTGCCCACTAAGTAGATAA 1519

RESULT 8

ACN89244

ID ACN89244 standard; DNA; 2027 BP.

XX ACN89244;

XX 02-DEC-2004 (first entry)

XX Breast cancer related marker, seq id 10394.

XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.

XX Homo sapiens.

XX US2003099974-A1.

XX 29-MAY-2003.

XX 18-JUL-2002; 2002US-00198846.

XX 18-JUL-2001; 2001US-0306220P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2003-787014/74.

XX Novel isolated polypeptide associated with breast cancer, useful for
XX detecting presence of polypeptide in sample, as a marker for breast
XX cancer.

XX Disclosure; SEQ ID NO 10394; 36pp; English.

XX The invention relates to an isolated polypeptide (I) associated with
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (SI). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the
CC invention may be described as cytostatic. The antibody is useful for
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC invention are useful in the detection of breast tumours. (I) is useful as
CC a marker for breast cancer and in breast cancer therapy. Sequences given
CC in records ACN78851-ACN92934 represent nucleic acid markers associated
CC with breast cancer. Note: The sequence listing does not form part of the
CC specification but may be obtained in electronic format from the USPTO web
CC site at seqdata.uspto.gov/sequence.html?DocID=20030099974

XX Sequence 2027 BP; 553 A; 551 C; 565 G; 354 T; 0 U; 4 Other;

Query Match 98.1%; Score 83.4; DB 11; Length 2027;

Best Local Similarity 98.8%; Pred. No. 2.6e-20;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60

Db 1282 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 1341

Qy 61 CTGGGTGCCCACTAAGTAGAAAA 85

Db 1342 CTGGGTGCCCACTAAGTAGATAA 1366

RESULT 9

ADJ56452

ID ADJ56452 standard; cDNA; 3674 BP.

XX ADJ56452;

XX 06-MAY-2004 (first entry)

XX Human cDNA differentially expressed in MYCN activated cells SeqID 258.

XX human; differential expression; transactivator; proto-oncogene;

XX neuroblastoma; small cell lung cancer; cytostatic; gene therapy; ss;

XX MYCN activated cell.

```
XX Homo sapiens.
OS
XX
XX US2003119009-A1.
PN
XX
XX 26-JUN-2003.
PD
XX
XX
XX 25-FEB-2002; 2002US-00084817.
PF
XX
XX 23-FEB-2001; 2001US-0270784P.
PR
XX (STUA/) STUART S G.
PA (NUCH/) NUCHTERN J G.
PA (PLON/) PLON S E.
PA (SHOH/) SHOHEIT J M.
XX
XX Stuart SG, Nuchtern JG, Plon SE, Shohet JM;
XX WPI; 2003-635698/60.
XX
XX New genes regulated by MYCN activation, useful in gene therapy,
XX particularly for treating a subject with e.g. neuroblastoma or other
XX cancers, or for diagnosing, staging or monitoring the treatment of the
XX cancer.
XX
XX Claim 1; SEQ ID NO 258; 27pp; English.
XX
XX This invention relates to novel isolated cDNAs that are differentially
XX expressed in MYCN activated cells. Specifically, it refers to
XX polynucleotide sequences that exhibit differential expression patterns in
XX cells activated by the transactivator MYCN, where MYCN is a proto-
XX oncogene that is amplified in neuroblastoma cells and is common in small
XX cell lung cancers. The present invention describes these cDNA molecules
XX as useful for in hybridisation assays to detect expression of nucleic
XX acids (or complementary nucleic acids) in a present in a given sample, as
XX well as for screening assays by identifying molecules or compounds that
XX specifically bind the cDNA as a ligand and modulate function or activity.
XX Accordingly, these compositions exhibit cytostatic activity and can also
XX be used for gene therapy purposes. This polynucleotide sequence is a cDNA
XX that is differentially expressed in MYCN activated cells, given in an
XX exemplification of the invention. NOTE: This sequence does not appear in
XX the printed specification but has been obtained in electronic format from
XX the US Patent Office at
XX ftp.seqdata.uspro.gov/sequence.html?DocID=20030119009.
XX
XX Sequence 3674 BP; 737 A; 1036 C; 1075 G; 826 T; 0 U; 0 Other;
XX
XX Query Match 98.1%; Score 83.4; DB 10; Length 3674;
XX Best Local Similarity 98.8%; Pred. No. 3.1e-20;
XX Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 1 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX Db 3501 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 3560
XX
XX Qy 61 CTGGTGCCCCCACTAAGTAGAAAA 85
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX Db 3561 CTGGTGCCCCCACTAAGTAGAATAA 3585
XX
XX RESULT 10
XX ABL62723
XX ID ABL62723 standard; DNA; 66109 BP.
XX
XX AC ABL62723;
XX
XX 15-MAY-2002 (first entry)
XX
XX Colon adenocarcinoma related gene sequence SEQ ID NO:1060.
XX
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
XX cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
```

```
KW gene; ds.
XX
XX Homo sapiens.
XX
XX WO200194629-A2.
PN
XX
XX 13-DEC-2001.
PD
XX
XX 30-MAY-2001; 2001WO-US010838.
PF
XX
XX 05-JUN-2000; 2000US-0209473P.
XX
XX 05-JUN-2000; 2000US-0209531P.
XX
XX 18-SEP-2000; 2000US-0233133P.
XX
XX 18-SEP-2000; 2000US-0233617P.
XX
XX 20-SEP-2000; 2000US-0234009P.
XX
XX 20-SEP-2000; 2000US-0234034P.
XX
XX 20-SEP-2000; 2000US-0234052P.
XX
XX 22-SEP-2000; 2000US-0234509P.
XX
XX 22-SEP-2000; 2000US-0234567P.
XX
XX 25-SEP-2000; 2000US-0234923P.
XX
XX 25-SEP-2000; 2000US-0234924P.
XX
XX 25-SEP-2000; 2000US-0235077P.
XX
XX 25-SEP-2000; 2000US-0235082P.
XX
XX 25-SEP-2000; 2000US-0235134P.
XX
XX 25-SEP-2000; 2000US-0235280P.
XX
XX 26-SEP-2000; 2000US-0235637P.
XX
XX 26-SEP-2000; 2000US-0235638P.
XX
XX 27-SEP-2000; 2000US-0235711P.
XX
XX 27-SEP-2000; 2000US-0235720P.
XX
XX 27-SEP-2000; 2000US-0235840P.
XX
XX 27-SEP-2000; 2000US-0235863P.
XX
XX 28-SEP-2000; 2000US-0236028P.
XX
XX 28-SEP-2000; 2000US-0236032P.
XX
XX 28-SEP-2000; 2000US-0236033P.
XX
XX 28-SEP-2000; 2000US-0236034P.
XX
XX 28-SEP-2000; 2000US-0236109P.
XX
XX 28-SEP-2000; 2000US-0236111P.
XX
XX 29-SEP-2000; 2000US-0236842P.
XX
XX 29-SEP-2000; 2000US-0236891P.
XX
XX 02-OCT-2000; 2000US-0237172P.
XX
XX 02-OCT-2000; 2000US-0237173P.
XX
XX 02-OCT-2000; 2000US-0237278P.
XX
XX 02-OCT-2000; 2000US-0237294P.
XX
XX 02-OCT-2000; 2000US-0237295P.
XX
XX 02-OCT-2000; 2000US-0237316P.
XX
XX 03-OCT-2000; 2000US-0237425P.
XX
XX 03-OCT-2000; 2000US-0237598P.
XX
XX 03-OCT-2000; 2000US-0237604P.
XX
XX 03-OCT-2000; 2000US-0237606P.
XX
XX 03-OCT-2000; 2000US-0237608P.
XX
XX 01-NOV-2000; 2000US-0244867P.
XX
XX 01-NOV-2000; 2000US-0245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, and determining a change
XX in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 1060; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
XX neoplastic agent. The method involves exposing cells to a chemical agent
XX to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
```


CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour

XX SQ Sequence 66109 BP; 17149 A; 14664 C; 14177 G; 20118 T; 0 U; 1 Other;

Query Match 98.1%; Score 83.4; DB 6; Length 66109;

Best Local Similarity 98.8%; Pred. No. 8.3e-20;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60

Db 37840 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 37899

Qy 61 CTGGTGCCCCCACTAAGTAGAATAA 85

Db 37900 CTGGTGCCCCCACTAAGTAGAATAA 37924

RESULT 11

ABN97273
ID ABN97273 standard; DNA; 66109 BP.

XX AC ABN97273;

DT 13-AUG-2002 (first entry)

DE Gene #3771 used to diagnose liver cancer.

XX Gens; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.

XX OS Homo sapiens.

XX WO200229103-A2.

XX PD 11-APR-2002.

XX PF 02-OCT-2001; 2001WO-US030589.

XX PR 02-OCT-2000; 2000US-0237054P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX WPI; 2002-426119/45.

XX Diagnosing and detecting the progression of liver cancer, hepatocellular
PT carcinoma or metastatic liver tumor in a patient, involves detecting the
PT level of expression of two or more genes in a liver tissue sample.

XX PS Claim 1; SEQ ID NO 3771; 298pp; English.

XX The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumor in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,

CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 66109 BP; 17149 A; 14664 C; 14177 G; 20118 T; 0 U; 1 Other;

Query Match 98.1%; Score 83.4; DB 6; Length 66109;

Best Local Similarity 98.8%; Pred. No. 8.3e-20;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 60

Db 37840 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTT 37899

Qy 61 CTGGTGCCCCCACTAAGTAGAATAA 85

Db 37900 CTGGTGCCCCCACTAAGTAGAATAA 37924

RESULT 12

ABK84699
ID ABK84699 standard; cDNA; 198285 BP.

XX AC ABK84699;

XX 14-AUG-2002 (first entry)

XX Human cDNA differentially expressed in granulocytic cells #1270.

XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.

XX OS Homo sapiens.

XX WO200228999-A2.

XX PD 11-APR-2002.

XX PF 03-OCT-2001; 2001WO-US030821.

XX PR 03-OCT-2000; 2000US-0237189P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

XX WPI; 2002-435328/46.

XX Detecting granulocyte activation by detecting differential expression of
PT genes associated with granulocyte activation, which serves as diagnostic
PT markers that is useful for monitoring disease states and drug toxicity.

XX PS Claim 1; SEQ ID NO 1270; 114pp; English.

XX The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing the
CC expression level to an expression level in an unactivated GC, where
CC differential expression of Gs is indicative of GCA. Also included are
CC modulating (M2) GA by contacting GC with an agent that alters the
CC expression of at least one gene in Gs; (2) screening (M3) for an agent
CC capable of modulating GCA or an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease using the gene expression
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a

CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from Gs, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating Gs; M3 is useful
CC for screening an agent capable of modulating GCA preferably in an
CC inflammation in a tissue; M4 is useful for detecting an inflammation
CC (especially chronic) in a tissue, an allergic response in a subject,
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, periodontal disease; also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 198285 BP; 55426 A; 41198 C; 41497 G; 60164 T; 0 U; 0 Other;
Query Match 98.1%; Score 83.4; DB 6; Length 198285;
Best Local Similarity 98.8%; Pred. No. 1.2e-19;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAGCATTTCCCCCAGGGAAGTTT 60
DB 113629 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAGCATTTCCCCCAGGGAAGTTT 113688

QY 61 CTGGTGCCCCCACTAAGTAGAAAAA 85
DB 113689 CTGGTGCCCCCACTAAGTAGAAAAA 113713

RESULT 13
ID ABN97319
AC ABN97319 standard; DNA; 198285 BP.
XX
AC ABN97319;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #3817 used to diagnose liver cancer.
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN W0200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US030589.
XX
PX 02-OCT-2000; 2000US-0237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI; 2002-426119/45.
XX
DR
XX Diagnosing and detecting the progression of liver cancer, hepatocellular
PT carcinoma or metastatic liver tumor in a patient, involves detecting the
PT level of expression of two or more genes in a liver tissue sample.
XX

PS Claim 1; SEQ ID NO 3817; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 198285 BP; 55426 A; 41198 C; 41497 G; 60164 T; 0 U; 0 Other;
Query Match 98.1%; Score 83.4; DB 6; Length 198285;
Best Local Similarity 98.8%; Pred. No. 1.2e-19;
Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAGCATTTCCCCCAGGGAAGTTT 60
DB 113629 CGGTGAAGTGCATCTTCTCGGTGTTCTCTATTGAAACAGCATTTCCCCCAGGGAAGTTT 113688

QY 61 CTGGTGCCCCCACTAAGTAGAAAAA 85
DB 113689 CTGGTGCCCCCACTAAGTAGAAAAA 113713

RESULT 14
ID ADR52987
XX
AC ADR52987 standard; DNA; 198285 BP.
XX
AC ADR52987;
XX
DT 18-NOV-2004 (first entry)
XX
DE Drug therapy altered expressed gene #338.
XX
DE
XX
KW drug activity monitoring; expression profile; gene expression;
KW peripheral blood sample; peripheral blood mononuclear cell; drug therapy;
KW CCI-779; immunosuppressant; rapamycin; mammalian target of rapamycin;
KW mTOR; ds.
XX
OS Homo sapiens.
XX
PN W02004072265-A2.
XX
PD 26-AUG-2004.
XX
PX 11-FEB-2004; 2004WO-US004118.
XX
PR 11-FEB-2003; 2003US-0446133P.
PR 03-APR-2003; 2003US-0459782P.
PR 23-JAN-2004; 2004US-0538246P.
XX
PA (AMHP) WYETH.
PA (BURC/) BURCZYNSKI M.
PA (TWIN/) TWINE N.
PA (DORN/) DORNER A J.
PA (TREP/) TREPICCHIO W L.
XX
PI Burczynski M, Twine N, Dornier AJ, Trepicchio WL;
XX WPI; 2004-642301/62.
XX
DR
XX Monitoring drug activities in vivo comprises comparing an expression
PT profile of a gene in a peripheral blood sample of a patient before and
PT after drug therapy.
XX

XX Disclosure; SEQ ID NO 338; 136pp; English.

XX

XX The invention relates to a method of monitoring drug activities in vivo

CC by comparing an expression profile of at least one gene in a peripheral

CC blood sample of a patient to a reference expression profile of the at

CC least one gene, where the at least one gene is differentially expressed

CC in peripheral blood mononuclear cells (PBMCs) of patients who have a non-

CC blood disease and are subjected to a drug therapy as compared to PBMCs

CC isolated from the patient before the drug therapy, and where the patient

CC has the non-blood disease and is being treated by the drug therapy. The

CC method, kit, and nucleic acid array are useful for monitoring drug

CC activities in vivo. The drug is especially CCI-779, an ester analogue of

CC the immunosuppressant rapamycin which is a potent inhibitor of the

CC mammalian target of rapamycin (mTOR). This sequence represents a gene

CC expressed in PBMC altered by the drug therapy. (Note: this sequence does

CC no form part of the printed specification but was obtained in electronic

CC format from WIPO at http://wipo.int/pub/published_pct_sequences/).

XX

XX Sequence 198285 BP; 55426 A; 41198 C; 41497 G; 60164 T; 0 U; 0 Other;

XX

Query Match 98.1%; Score 83.4; DB 13; Length 198285;

Best Local Similarity 98.8%; Pred. No. 1.2e-19;

Matches 84; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAACAAGCATTTCCTCCCGGGAAGTTT 60

DB 113629 CGGTGAAGTGCATCTTCGCGTGTCTCTATTGAACAAGCATTTCCTCCCGGGAAGTTT 113688

QY 61 CTGGTGCCCCCACTAAGTAGAAAA 85

DB 113689 CTGGTGCCCCCACTAAGTAGATAA 113713

RESULT 15

AAC89745

ID AAC89745 standard; cDNA; 108 BP.

XX AAC89745;

AC AAC89745;

XX 12-MAR-2001 (first entry)

XX Human gastrointestinal inflammation-related cDNA, SEQ ID NO: 14.

XX Human; cytostatic; immunomodulator; immunostimulant; vulnerary;

KW anti-inflammatory; neuroprotective; antibacterial; gene therapy;

KW gastrointestinal inflammation; immune system disorder; genetic disorder;

KW cancer; autoimmune disorder; infection; wound healing; ss.

XX

OS Homo sapiens.

XX

XX WO200073324-A2.

XX

XX 07-DEC-2000.

XX

XX 01-JUN-2000; 2000WO-US015191.

XX

XX 01-JUN-1999; 99US-0137058P.

XX

PA (DIGI-) DIGITAL GENE TECHNOLOGIES INC.

XX

XX Youakim A, Dubose RF, Sims JE, Pribyl TM, Hillbush BS, Hasel KW;

PI "pI; 2001-061508/07.

CC which are useful in gene therapy, and for diagnosing a pathological
 CC condition or a susceptibility to it. In particular, the polynucleotides
 CC are useful for modulating gene expression in gastrointestinal
 CC inflammation. The polynucleotides are useful for chromosome
 CC identification, controlling gene expression through triple helix
 CC formation or antisense DNA or RNA, or identifying individuals from minute
 CC biological samples using DNA-based identification techniques. The
 CC polynucleotides can also be used as an alternative to restriction
 CC fragment length polymorphism (RFLP), by determining the actual base-by-
 CC base DNA sequences of selected portions of an individual's genome. The
 CC polynucleotides may also be used as molecular weight markers on Southern
 CC gels, as diagnostic probes for the presence of a specific mRNA, as a
 CC probe to subtract-out known sequences in the process of discovering novel
 CC polynucleotides, or as an antigen to elicit an immune response. The
 CC polypeptides are useful in diagnostic procedures to detect a disorder.
 CC The polynucleotides and polypeptides are useful for preventing, treating
 CC or ameliorating immune system disorders, genetic disorders, cancers, some
 CC autoimmune disorders, or infections. The polynucleotides and polypeptides
 CC are also useful for differentiating, proliferating or attracting cells,
 CC leading to the regeneration of tissues, especially in wounds or burns.
 CC The polypeptides and polynucleotides may also be used as a food additive
 CC or preservative

xx
 SQ Sequence 108 BP; 35 A; 20 C; 24 G; 29 T; 0 U; 0 Other;

Query Match	96.2%;	Score 81.8;	DB 4;	Length 108;
Best Local Similarity	97.6%;	Pred. No.	3.7e-20;	
Matches	83;	Conservative	0;	Mismatches 2;
Indels	0;	Gaps	0;	

Qy 1 CGGTGAAGTGCATCTTCGCGTGTTCTCTATTGGACAACAGCATTTCCCCCAGGGAAGTTT 60
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 1 CGGTGAAGTGCATCTTCGCGTGTTCTCTATTGGACAACAGCATTTCCCCCAGGGAAGTTT 60
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 61 CTGGTGTCGCCCACTAAGTAGAAAAA 85
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 61 CTGGTGTCGCCCACTAAGTAGAAATAA 85
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Search completed: December 7, 2005, 00:03:17
 Job time : 136.627 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2005, 10:04:41 ; Search time 37.0997 Seconds
(without alignments)
4072.612 Million cell updates/sec

Title: US-09-980-046B-10
Perfect score: 85
Sequence: 1 cggggaagtcacatctctgc.....tgccccactaagtagaaaaa 85

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
7: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
8: /cgn2_6/ptodata/1/ina/BACKFILES.seq.*
9: /cgn2_6/ptodata/1/ina/BACKFILES.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	28.2	33.2	21513	3	Sequence 16695, A
C 2	28.2	33.2	21513	3	Sequence 16696, A
C 3	28.2	33.2	21513	3	Sequence 16697, A
C 4	26.4	31.1	45138	3	Sequence 13027, A
C 5	26	30.6	856	3	Sequence 3, Appli
C 6	26	30.6	10103	3	Sequence 11967, A
C 7	26	30.6	10103	3	Sequence 13245, A
C 8	25.6	30.1	323	3	Sequence 30159, A
C 9	25.6	30.1	2658	3	Sequence 324, App
10	25.6	30.1	3601	3	Sequence 2, Appli
C 11	25.4	29.9	818128	3	Sequence 3880, Ap
C 12	25.4	29.9	818128	3	Sequence 14546, A
C 13	25.4	29.9	818128	3	Sequence 14547, A
C 14	25.4	29.9	818128	3	Sequence 14548, A
C 15	25.4	29.9	818128	3	Sequence 14549, A
C 16	25.4	29.9	818128	3	Sequence 14550, A
C 17	25.4	29.9	818128	3	Sequence 14551, A
C 18	25.4	29.9	818128	3	Sequence 14552, A
C 19	25.4	29.9	818128	3	Sequence 14553, A
C 20	25.4	29.9	818128	3	Sequence 14554, A
C 21	25.4	29.9	818128	3	Sequence 14555, A
C 22	25.4	29.9	818128	3	Sequence 14556, A
C 23	25.4	29.9	818128	3	Sequence 14557, A
C 24	25.4	29.9	818128	3	Sequence 14558, A

ALIGNMENTS

RESULT 1

US-09-949-016-16695/c
; Sequence 16695, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16695
; LENGTH: 21513
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16695

Query Match 33.2%; Score 28.2; DB 3; Length 21513;
Best Local Similarity 64.6%; Pred. No. 3.1; Mismatches 23; Indels 0; Gaps 0;
Matches 42; Conservative 0;

Qy 9 TGCATCTCTCGGTTCTCTATTGAACAGCATTTCCCGGAGGTTCTCGGGTC 68
Db 7635 TCCATAGCTGCCCTTTCTTTTAAAGCTTGAAGCTCTGAGCTCTGAGGTTTCAGGGTTC 7576
Qy 69 CCCAC 73
Db 7575 CCCAC 7571

RESULT 2

US-09-949-016-16696/c
; Sequence 16696, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16696
; LENGTH: 21513
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16696

Query Match 33.2%; Score 28.2; DB 3; Length 21513;
Best Local Similarity 64.6%; Pred. No. 3.1;
Matches 42; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 9 TGCATCTCTCGGTGTTCTCTATTGAACAAGCATTTCCCCAGGGAAGTTTCTGGGTGC 68
Db 7635 TCCATAGCTGCCTTTTCTTTAAATTTCTAAAGCTTTGAGCCTCTGGATTTTTCAGGGTTC 7576

QY 69 CCCAC 73
Db 7575 CCCAC 7571

RESULT 3
US-09-949-016-16697/c
; Sequence 16697, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16697
; LENGTH: 21513
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16697

Query Match 33.2%; Score 28.2; DB 3; Length 21513;
Best Local Similarity 64.6%; Pred. No. 3.1;
Matches 42; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 9 TGCATCTCTCGGTGTTCTCTATTGAACAAGCATTTCCCCAGGGAAGTTTCTGGGTGC 68
Db 7635 TCCATAGCTGCCTTTTCTTTAAATTTCTAAAGCTTTGAGCCTCTGGATTTTTCAGGGTTC 7576

QY 69 CCCAC 73
Db 7575 CCCAC 7571

RESULT 4
US-09-949-016-13027
; Sequence 13027, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13027
; LENGTH: 45138
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(45138)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13027

Query Match 31.1%; Score 26.4; DB 3; Length 45138;
Best Local Similarity 69.2%; Pred. No. 20;
Matches 36; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 23 GTTCTCTATTGAACAAGCATTTCCCCAGGGAAGTTTCTGGGTGCCCACT 74
Db 15383 GTTATTTATTAAACAAGAATTTCTGACAAAGAAGTTTGTATGTCCTCTCT 15434

RESULT 5
US-09-617-543-3/c
; Sequence 3, Application US/09617543
; Patent No. 6849776
; GENERAL INFORMATION:
; APPLICANT: KUVSHINOV, Viktor
; APPLICANT: KOIVU, Kimmo
; APPLICANT: KANERVA, Anne
; APPLICANT: PEHY, Eija
; TITLE OF INVENTION: MOLECULAR CONTROL OF TRANSGENE SEGREGATION AND ESCAPE
; BY RECOVERABLE BLOCK OF FUNCTION (RBF) SYSTEM
; FILE REFERENCE: KUVSHINOV=1
; CURRENT APPLICATION NUMBER: US/09/617,543
; CURRENT FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 856
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Vigna mungo
; (SH-EP promoter), Bacillus amyloliquefaciens
; OTHER INFORMATION: (barnase gene), Escherichia coli (uidA gene)
; NAME/KEY: promoter
; LOCATION: (1)..(352)
; OTHER INFORMATION: Modified SH-EP promoter.
; NAME/KEY: misc.feature
; LOCATION: (353)..(404)
; OTHER INFORMATION: 5' end exon of modified uidA gene.
; NAME/KEY: misc.feature
; LOCATION: (443)..(778)
; OTHER INFORMATION: Synthetic coding sequence of barnase gene.
; NAME/KEY: intron
; LOCATION: (405)..(830)
; OTHER INFORMATION: Intron of uidA gene.
; NAME/KEY: misc.feature
; LOCATION: (831)..(856)
; OTHER INFORMATION: Exon of modified uidA gene.
US-09-617-543-3

QY	71	CACTAAGTAGAA	82
Db	24	TACTAAGACAA	13

RESULT 9
US-09-248-796A-324
; Sequence 324, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 324
; LENGTH: 2658
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-324

;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 14547
;; LENGTH: 818128
;; TYPE: DNA
;; ORGANISM: Human
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)...(818128)
;; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14547

Query Match 29.9%; Score 25.4; DB 3; Length 818128;
Best Local Similarity 58.7%; Pred. No. 1.4e+02;
Matches 44; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 3 GTGAAGTCATCTTCGCTGTTCTCTATTGAAACAAGCATTTCCCCAGGAAAGTTTCT 62
Db 339125 GTGAATACCTTTTTCCTTGCTTCATTAAATATATATGACCCAGTTTGTGTTTT 339184

Qy 63 GGGTGCCCCACTAAG 77
Db 339185 GTTTTCCTTTTAG 339199

RESULT 14
US-09-949-016-14548
;; Sequence 14548, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; CURRENT FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 14548
;; LENGTH: 818128
;; TYPE: DNA
;; ORGANISM: Human
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)...(818128)
;; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14548

Query Match 29.9%; Score 25.4; DB 3; Length 818128;
Best Local Similarity 58.7%; Pred. No. 1.4e+02;
Matches 44; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 3 GTGAAGTCATCTTCGCTGTTCTCTATTGAAACAAGCATTTCCCCAGGAAAGTTTCT 62
Db 339125 GTGAATACCTTTTTCCTTGCTTCATTAAATATATATGACCCAGTTTGTGTTTT 339184

Qy 63 GGGTGCCCCACTAAG 77
Db 339185 GTTTTCCTTTTAG 339199

RESULT 15

US-09-949-016-14549
;; Sequence 14549, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; CURRENT FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 14549
;; LENGTH: 818128
;; TYPE: DNA
;; ORGANISM: Human
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)...(818128)
;; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14549

Query Match 29.9%; Score 25.4; DB 3; Length 818128;
Best Local Similarity 58.7%; Pred. No. 1.4e+02;
Matches 44; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 3 GTGAAGTCATCTTCGCTGTTCTCTATTGAAACAAGCATTTCCCCAGGAAAGTTTCT 62
Db 339125 GTGAATACCTTTTTCCTTGCTTCATTAAATATATATGACCCAGTTTGTGTTTT 339184

Qy 63 GGGTGCCCCACTAAG 77
Db 339185 GTTTTCCTTTTAG 339199

Search completed: December 6, 2005, 11:47:38
Job time : 47.0997 secs

This Page Blank (uspto)

; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63714
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Bovine 19866881014087
US-10-750-185-63714

Query Match 31.1%; Score 26.4; DB 6; Length 1383;
Best Local Similarity 59.2%; Pred. No. 0.9;
Matches 45; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 4 TGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTTCTG 63
Db 303 TGCAGTCGAGAGACAGCGGTTTCATTCCTCGTCAGGAAGTCCCCAGAGAAGGGAGTG 244

Qy 64 GGTCCCCCACTAAGTA 79
Db 243 GCTCCCCCACTCAGTA 228

RESULT 3
US-10-750-185-46975
; Sequence 46975, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46975
; LENGTH: 3331
; TYPE: DNA
; ORGANISM: Bovine 19866881406138
US-10-750-185-46975

Query Match 30.1%; Score 25.6; DB 6; Length 3331;
Best Local Similarity 70.8%; Pred. No. 2.6;
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CGGTGAAGTGCATCTTCTGCGTGTCTCTATTGAAACAAGCATTTCCC 48
Db 1778 CTGTAAGAGAAAGAACTTCTCCTCATCTACTATTGTAATAGTATTTCCC 1825

RESULT 4
US-10-750-185-44982/c
; Sequence 44982, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.

; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44982
; LENGTH: 1846
; TYPE: DNA
; ORGANISM: Bovine 19866880725399
US-10-750-185-44982

Query Match 29.9%; Score 25.4; DB 6; Length 1846;
Best Local Similarity 68.6%; Pred. No. 2.5;
Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 33 TGAACAAGCATTTCCCCAGGGAAGTTTCTGGGTGCCCACTAAGTAGAAA 83
Db 1347 TGAACAACACTATGCCATAAGGAAGGCTGCTGGATGCTCTACTATTGAGAAA 1297

RESULT 5
US-10-750-185-58054/c
; Sequence 58054, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58054
; LENGTH: 2786
; TYPE: DNA
; ORGANISM: Bovine 19866880580668
US-10-750-185-58054

Query Match 29.9%; Score 25.4; DB 6; Length 2786;
Best Local Similarity 61.2%; Pred. No. 2.9;
Matches 41; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 19 GCGTGTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTTCTGGGTGCCCACTAAGT 78
Db 2267 GGCATCTCTTCATATGTACAAAGATTCTCTAGAAAAAATATTAGGAACCCCTCTAAGT 2208

Qy 79 AGAAAAA 85
Db 2207 AAAATAA 2201

RESULT 6
US-10-750-185-31717/c
; Sequence 31717, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:

```
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 31717
; LENGTH: 1846
; TYPE: DNA
; ORGANISM: Bovine 19866880465309
US-10-750-185-31717

Query Match          29.6%; Score 25.2; DB 6; Length 1846;
Best Local Similarity 62.9%; Pred. No. 2.9;
Matches 39; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 23 GTTCTCTATTGAACAAGCATTTCCCCAGGGAAGTTTCTGGTGCCCACTTAAGTAGAA 82
Db 1680 GATTCTTATGTCACAAACCTTCAAGATGATGTTGGGGGCATCCACACGAGAAA 1621

QY 83 AA 84
Db 1620 AA 1619
```

```
RESULT 7
US-10-750-185-36838/c
; Sequence 36838, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 36838
; LENGTH: 2518
; TYPE: DNA
; ORGANISM: Bovine 19866880875187
US-10-750-185-36838

Query Match          29.4%; Score 25; DB 6; Length 2518;
Best Local Similarity 58.9%; Pred. No. 4;
Matches 43; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
```

```
QY 2 GGTGAAGTCATCTTCTGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTTC 61
Db 2434 GCTGCAGTGTAGATTCTGGGTCTCCTGTAGGAGACCCAGGTTTGATCCCTGGGTCGAGTC 2375

QY 62 TGGTGCCCACT 74
Db 2374 AGGAAGATCCCT 2362
```

```
RESULT 8
US-10-750-185-44918
; Sequence 44918, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 44918
; LENGTH: 4341
; TYPE: DNA
; ORGANISM: Bovine 19866880547884
US-10-750-185-44918

Query Match          29.2%; Score 24.8; DB 6; Length 4341;
Best Local Similarity 60.3%; Pred. No. 5.9;
Matches 41; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 12 ATCTTCTGCGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTTCTGGGTGCCCC 71
Db 3179 ATCTAATGTCATTCATGATTTTAAAGCAACATTCACGGGGAATTCATGCTGGTGCCC 3238

QY 72 ACTAAGTA 79
Db 3239 AGTGCCTA 3246
```

```
RESULT 9
US-10-750-185-43944/c
; Sequence 43944, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 43944
; LENGTH: 1364
; TYPE: DNA
; ORGANISM: Bovine 19866881010494
US-10-750-185-43944

Query Match          28.9%; Score 24.6; DB 6; Length 1364;
Best Local Similarity 57.0%; Pred. No. 4.4;
Matches 45; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
```

```
QY 6 AAGTCATCTTCTGGTGTTCTCTATTGAAACAAGCATTTCCCCAGGGAAGTTTCTGGG 65
Db 851 AATTGGAACCTATTTGTAGGTTCTGTTGTGGAACATTTAGTTCTGTTGTGGA 792
```

QY 66 TGGCCCACTAAGTAGAAAA 84
 |||||||
Db 791 AACCCCACATCATAGAAAA 773

```

RESULT 10
US-10-750-185-46482
; Sequence 46482, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MW1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46482
; LENGTH: 1535
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-46482

```

	Query Match	28.5%	Score 24.2	DB 6	Length 1535
	Best Local Similarity	57.1%	Fred. No. 6.6		
	Matches 44	Conservative 0	Mismatches 33	Indels 0	Gaps 0
Qy	2	GGTGAAGTGCATCTCTCGGTGTTCTATTATGAACACAGCATTTCCCCCGAGGAAGTTTC	61		
Db	153	GGTGAGGCCACCTCTGCTCTGCTTTTTTTTAAATCCCTTTTACCCCTTTACTTGA	212		
Qy	62	TGGGTGCCCCCACTAAGT	78		
Db	213	TACCTCCCTCCCTGTCACT	229		

```

RESULT 11
US-10-750-185-47332/c
; Sequence 47332, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47332
; LENGTH: 905
; TYPE: DNA
; ORGANISM: Bovine 19866881656724
US-10-750-185-47332

```

Query Match	28.2%	Score 24;	DB 6;	Length 905;
Best Local Similarity	64.3%	Pred. NO. 6.4;		
Matches 36;	Conservative	0;	Mismatches 20;	Indels 0;
Gaps	0;			

QY 29 TATTTGAACAAGCATTTCCCCAGGGAAGTTTCTGGTGCCCACTAAGTAGAAAA 84

Db 127 TATTTTAAATAGCAATACCCCATCCACGTTTCTTAGCTGTTAAGTATGCAGAAAA 72

```

RESULT 12
US-10-750-185-37876
; Sequence 37876, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 37876
; LENGTH: 1795
; TYPE: DNA
; ORGANISM: Bovine 19866880803754
US-10-750-185-37876

```

Query March	28.2%	Score 24;	DB 6;	Length 1795;
Best Local Similarity	60.9%	Pred. NO. 8.4;		
Matches 39;	Conservative 0;	Mismatches 25;	Indels 0;	Gaps 0;
QY	13	TCMTCTCGTGTTCTCTATTGTGAACAAGCATTTCCCCCAGGGAAGTTTCTGGGTGCCCA	72	
Db	1370	TATTATTCTCCCTCCCTCTCTCAAAAACCCCTGCAAGGAAGTTTCTGGATGCCAA	1429	
QY	73	CTAA	76	
Db	1430	ACAA	1433	

```

RESULT 13
US-10-750-185-1830
; Sequence 1830. Application US/10750185
; Publication No. US2005026030A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 1830
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT21454
US-10-750-185-1830

```

Query Match 28.0%; Score 23.8; DB 6; Length 600;
Best Local Similarity 55.4%; Pred. No. 6.5;

	Matches	46;	Conservative	0;	Mismatches	37;	Indels	0;	Gaps	0;
Qy	2	GGTGAAGTCATCTTCTCGTGTTCTCTATTGAAACAAAGCATTTCCCCCAGGGAAAGTTTC	61							
Dd	187	GCATTATGFGGTTCCTTTGTTGTTGTTTTTTCGGTCATGCCCTATAGCATATAGCAAGATCT	246							
Qy	62	TGGGTGCCCCCACTAAGTAGAAAA	84							
Dd	247	TAGTTCCCCAACCCAGGGAGTGAA	269							

RESULT 14

```

RESUMI 14
US-10-750-185-1960
; Sequence 1960, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1960
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT21455
US-10-750-185-1960

```

RESULT 15

```

RES001 13
US-10-750-185-59439
; Sequence 59439, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: PANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59439
; LENGTH: 1166
; TYPE: DNA

```

Search completed: December 6, 2005, 23:16:55
Job time : 58.0228 secs

This Page Blank (uspto)